IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Applicants: Maurice M. MOLONEY, *et al.*

Title: **METHOD FOR PRODUCING AND CLEAVING A FUSION PROTEIN WITH AN N-TERMINAL CHYMOSIN PRO-PEPTIDE**

Appl. No.: 09/402,488 Filing Date: 2/16/2000

Examiner: David J. Steadman Art Unit: 1656

Confirmation Number: 6010

APPEAL BRIEF

Mail Stop Appeal Brief-Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

The is an appeal under 35 U.S.C. § 134 of the Examiner's final decision in the final Office action dated July 12, 2006, rejecting claims 1, 4-10, 12-16, 18-19, and 50-51. A Notice of Panel Decision from Pre-Appeal Brief Review determining that the appeal should proceed was issued January 19, 2007.

This Appeal Brief is being filed together with one credit card payment form in the amount of $950, covering the appeal fee ($500), and the extension fee ($450). By virtue of the concurrently filed Petition for Extension of Time and payment of the prescribed fee, the Appeal Brief is timely filed. If this payment is deemed to be insufficient, authorization is hereby given to charge any deficiency (or credit any overpayment) to deposit account no. 19-0741.

TABLE OF CONTENTS

TABLE OF AUTHORITIES

FEDERAL CASES

FEDERAL STATUTES

REGULATIONS

I. REAL PARTY IN INTEREST

The real party in interest is SemBioSys Genetics Inc., Canada, which is the assignee of each inventor's entire interest.

II. RELATED APPEALS AND INTERFERENCES

No related appeals or interferences are pending.

III. STATUS OF CLAIMS

Claims 1, 4-10, 12-16, 18-19, and 50-51 are finally rejected and are the subject of this appeal.

No claims are allowed.

No claims are subject to objection.

IV. STATUS OF AMENDMENTS

No post-final amendments or submissions were denied entry into the application. In the Advisory Action mailed November 7, 2006, at page 1, Examiner Steadman indicated entry of the amendment filed October 11, 2006.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The claimed invention relates to the discovery that a recombinant polypeptide of interest can be obtained by a method that comprises obtaining a fusion protein comprising a full-length chymosin pro-peptide linked to the polypeptide of interest and contacting the fusion protein with a mature form of an autocatalytically maturing aspartic protease, whereby the chymosin pro-peptide is cleaved from the fusion protein to release the recombinant polypeptide. *See* Specification, page 1, lines 1-7.

As explained at pages 1-2 of the specification, the production of a recombinant polypeptide of interest as a fusion protein overcomes a number of problems that occur when the recombinant polypeptide is produced in its active form directly. For example, overproduced polypeptides can aggregate in a host cell in insoluble fractions, known as inclusion bodies, and obtaining usable protein from this insoluble material involves often slow and complex refolding methods which make protein purification difficult. Additionally, when the recombinantly produced polypeptides are present in soluble form in the cytoplasm, they can be degraded by host-specific enzymes, thus reducing the amount of active protein that can be recovered. Linking the protein of interest to a fusion partner has been found to limit these problems. *See* Specification, pages 1-2.

Typically, the preparation of a polypeptide of interest from a fusion protein involves preparing expression vectors comprising hybrid genes (nucleic acid molecules) that encode the polypeptide of interest linked to a fusion partner. Production of the polypeptide of interest thus involves introducing the hybrid gene into a host cell that will express (produce) the fusion protein, and then cleaving the fusion partner from the fusion protein to obtain the polypeptide of interest. *See* Specification, page 1. The cleavage may be effected by using enzymes or chemical reagents that hydrolyze peptide bonds, and enzymes known as "proteolytic enzymes" have been found to be particularly well suited for the cleavage of fusion proteins. *See* Specification, page 2, lines 10-15.

Prior to the present invention, common problems with fusion protein methodology included inefficient cleavage reactions, which result in low protein purification efficiency, and lack of cleavage reagent specificity, such that the fusion protein is cleaved at different sites, resulting in product loss and contamination with other polypeptides. *See* Specification, page 2, lines 28-35.

The present invention solves these problems by using a particular fusion partner, a chymosin pro-peptide, and a particular type of cleavage reagent, an autocatalytically maturing aspartic protease, which results in efficient and specific cleavage of the chymosin pro-peptide from the fusion protein to yield the recombinant polypeptide of interest. Chymosin, like other zymogens, is an enzyme that is synthesized as an inactive precursor ("zymogen") *in vivo*. *See* Specification, page 3, lines 1-3. Under appropriate conditions, zymogens are activated to form the mature active protein in a process involving the cleavage of an amino-terminal peptide which is referred to as the "pro-peptide," "pro-region" or "pro-sequence." *Id.* at lines 2-5. It is the chymosin pro-peptide sequence that is used in accordance with the invention as the fusion partner for the recombinant polypeptide of interest. *Id.* at lines 20-24.

Some zymogens are activated only in the presence of an additional proteolytic enzyme, while others are activated without an additional enzymatic catalyst. *Id.* at lines 5-8. It is enzymes of the latter category, referred to as "autocatalytically maturing" zymogens, that are used in the context of the present invention. *Id.* at lines 8-11. Examples of autocatalytically maturing zymogens include aspartic proteases, such as chymosin, pepsin, cathepsin, and yeast proteinase. *Id.* at page 10, lines 20-24.

The claims on appeal include two independent claims, claim 1 and claim 51. In general terms, the claimed methods involve obtaining a fusion protein comprising the recombinant polypeptide of interest and a chymosin pro-peptide and contacting the fusion protein with a mature form of an autocatalyically maturing aspartic protease. The aspartic protease cleaves the chymosin pro-peptide from the fusion protein, thereby releasing the recombinant polypeptide of interest. *See* Specification page 1, lines 3-7, and page 3, lines 20-27.

The method recited in claim 1 comprises (a) transforming a non-human host cell with an expression vector for the fusion protein; (b) growing the cell to produce the fusion protein; (c) obtaining the fusion protein from the host cell, and (d) contacting the fusion protein with a mature form of an autocatalytically maturing aspartic protease. *Id.* at page 3, line 35-page 4, line 18. As defined in claim 1 and explained above, the autocatalytically maturing aspartic protease is capable of cleaving the chymosin pro-peptide; thus, contacting it with the fusion protein results in cleavage of the chymosin pro-peptide from the fusion protein and release of the recombinant polypeptide. Specification, page 4, lines 17-18.

As reflected in claim 1, the expression vector comprises (1) a nucleic acid sequence capable of regulating transcription in a host cell, that is operatively linked to (2) a chimeric nucleic acid sequence that encodes a fusion protein, that is operatively linked to (3) a nucleic acid sequence encoding a termination region that is functional in said host cell. *Id.* at page 4, lines 1-9. The chimeric nucleic acid sequence comprises (a) a nucleic acid sequence encoding a full-length chymosin pro-peptide, that is linked in reading frame to (b) a nucleic acid sequence that is heterologous to the pro-peptide and that encodes the recombinant polypeptide. *Id.* The heterologous nucleic acid sequence is located immediately downstream of the nucleic acid sequence encoding the chymosin pro-peptide. *Id.*

Independent claim 51 is largely identical to claim 1, but specifies that the host cell is a bacterial cell, yeast cell, or plant cell, *see* Specification, page 7, lines 1-4, and that contacting the fusion protein with the aspartic protease is effected *in vivo*, *see* Specification, page 12, line 32.

Thus, the present invention harnesses the efficiency and specificity of autocatalytically maturing aspartic proteases for the chymosin pro-peptide sequence, and uses them in methods for obtaining recombinant polypeptides of interest. The examples reported in the specification demonstrate the efficacy of this methodology for producing model proteins such as hirudin and carp growth hormone in host cells such as *E. coli*. *See* Example 1, page 14, line 25-page 16, line 8; Example 2, page 16, line 10-page 17, line 15.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The grounds of rejection for review are as follows:

A. The rejection of claims 1, 4-10, 12-16, 18-19, and 50 under 35 U.S.C. § 112, first paragraph (written description).

B. The rejection of claims 1, 4-10, 12-16, 18-19, and 50-51 under 35 U.S.C. § 112, first paragraph (enablement), as applied to (i) claims 1, 4-10, 12-16, 18-19, and 50 and (ii) claim 51.

C. The rejection of claims 1, 4, 6-9, 13, 15, 19 and 51 under 35 U.S.C. § 103(a) over Ward (U.S. Patent No. 6,265,204) in view of Walsh (*J. Biotech* 45: 235-241 (1996)) and Yonezawa (*Int J Pept Protein Res* 47:56-61 (1996)).

D. The rejection of claim 5 under 35 U.S.C. § 103(a) over Ward (U.S. Patent No. 6,265,204) in view of Walsh (*J. Biotech* 45: 235-241 (1996)) and Yonezawa (*Int J Pept Protein Res* 47:56-61 (1996)), further in view of Fine (*Gen Comp Endocrinol.* 89:51-61 (1993)).

E. The rejection of claims 14 and 50 under 35 U.S.C. § 103(a) over Ward (U.S. Patent No. 6,265,204) in view of Walsh (*J. Biotech* 45: 235-241 (1996)) and Yonezawa (*Int J Pept Protein Res* 47:56-61 (1996)), further in view of Dunn ("Aspartic Proteases," ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY, Volume 362, Plenum Press, NY, 1995, pp. 1-9).

VII. ARGUMENT

A. Claims 1, 4-10, 12-16, 18-19, and 50 are supported by a written description as required by 35 U.S.C. § 112, first paragraph.

The Examiner rejected claims 1, 4-10, 12-16, 18-19, and 50 for alleged lack of written description with respect to the recitation of the transformation of a "non-human host cell." *See* Final Office Action of July 12, 2006, item 7, pages 3-4. This rejection is based on clear error in the Examiner's application of 35 U.S.C. § 112 and in his misinterpretation of the relevant MPEP provisions.

The written description requirement is embodied in the first paragraph of 35 U.S.C. § 112, which requires that "the specification shall contain a written description of the invention." 35 U.S.C. § 112 (2006). The written description requirement ensures "that patentees adequately describe their inventions in their patent specifications in exchange for the right to exclude others from practicing the invention for the duration of the patent's term." MPEP § 2163. As explained in MPEP § 2163, "to satisfy the written description requirement, a patent specification must describe the claimed invention in sufficient detail that one skilled in the art can reasonably conclude that the inventor had possession of the claimed invention" MPEP § 2163 (citing *Moba, B. V. v. Diamond Automation, Inc.*, 325 F.3d 1306, 1319, 66 U.S.P.Q.2d 1429, 1438 (Fed. Cir. 2003); *Vas-Cath, Inc. v. Mahurkar*, 935 F.2d 1563, 19 U.S.P.Q.2d 1116 (Fed. Cir. 1991)).

The "fundamental factual inquiry" for written description "is whether the specification conveys with reasonable clarity to those skilled in the art that, as of the filing date sought, applicant was in possession of the invention as now claimed." MPEP § 2163.02 (citing *Vas-Cath, Inc. v. Mahurkar*, 935 F.2d 1555, 1563-1564, 19 U.S.P.Q.2d 1111, 1117 (Fed. Cir. 1991)). "Possession may be shown in a variety of ways including description of an actual reduction to practice...or by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention." MPEP § 2163.02. Importantly, "[t]he subject matter of the claim need not be described literally (i.e., using the same terms or *in haec verba*) in order for the disclosure to satisfy the description requirement." MPEP § 2163.02.

The instant specification provides ample support for the transformation of "non-human" host cells, and clearly conveys to one of ordinary skill in the art that Appellant had possession of the claimed subject matter at the time of filing.

1. The specification describes the transformation of non-human host cells.

The rejected claims are directed to methods of preparing a recombinant polypeptide that involves transforming a non-human host cell with an expression vector. The specification describes the transformation of non-human host cells in the context of the recited methods. For example, the specification at page 7, lines 3-4, teaches that "fusion proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using, for example baculovirus), yeast cells, plant cells or mammalian cells." The specification additionally teaches that the protein "may also be produced in an edible food source, such as animal milk, or in an edible crop." Specification, page 13, lines 27-29.

These teachings, particularly when read in view of the disclosure as a whole, plainly convey to the skilled artisan Appellant's possession of the invention with regard to the transformation of non-human host cells. Indeed, because aspects of the invention relate to "methods for recovering recombinantly produced polypeptides" from sources such as "animal milk" and "an edible crop," it is difficult to imagine how the skilled artisan could reach a contrary conclusion. *See*, specification, pg. 1, lines 3-4, & Abstract,

Because the specification satisfies the written description requirement of § 112 with respect to the transformation of "non-human" host cells, the written description rejection of claims 1, 4-10, 12-16, 18-19, and 50 is improper, and should be reversed.

2. The written description rejection is improperly based on an *in haec verba* requirement.

The Examiner has admitted that the specification teaches numerous non-human host cells, such as bacterial cells, insect cells, yeast cells, and plant cells. *See* Advisory Action of

November 7, 2006, page 3. Nevertheless, he cites MPEP § 2173.05(i) for the proposition that "[a]ny negative limitation . . . must have basis in the original disclosure," apparently reading this provision as imposing a requirement for support *in haec verba* for a negative proviso. *Id.* at pages 3-4. Yet the MPEP imposes no such requirement, which would contravene settled law on the written description requirement.

As noted in MPEP § 2173.05(i), "there is nothing inherently ambiguous or uncertain about a negative limitation." The MPEP does state that "[a]ny negative limitation . . . must have *basis* in the original disclosure," but this does not require support *in haec verba*. To the contrary, the MPEP explains that "a lack of literal basis in the specification for a negative limitation may not be sufficient to establish a *prima facie* case for lack of descriptive support," and cites MPEP § 2163 for further information. The latter provisions plainly state that "the subject matter of the claim need not be described literally (*i.e.*, using the same terms or *in haec verba*) in order for the disclosure to satisfy the description requirement." Instead, it explains, the "fundamental factual inquiry" for written description "is whether the specification conveys with reasonable clarity to those skilled in the art that, as of the filing date sought, applicant was in possession of the invention as now claimed."

MPEP § 2163.04 sheds further light on this issue, requiring that a rejection for lack of written description include "reasons why a person skilled in the art at the time the application was filed would not have recognized that the inventor was in possession of the invention as claimed in view of the disclosure of the application as filed." The rejection here fails to include such reasoning.

As demonstrated above, the specification satisfies the fundamental factual written description inquiry, and demonstrates Appellant's possession of the claimed invention with regard to the transformation of non-human host cells. Accordingly, the written description rejection should be reversed.

B. Claims 1, 4-10, 12-16, 18-19, and 50-51 are supported by an enabling disclosure as required by 35 U.S.C. § 112.

The Examiner rejected claims 1, 4-10, 12-16, 18-19, and 50-51 for alleged lack of enablement with respect to the recombinant production of proteins in animals and plants. This rejection is based on clear error in the Examiner's application of the enablement requirements of 35 U.S.C. § 112. Although the Examiner did not distinguish claim 51 in this rejection, Appellant argues the patentability of claim 51 separately in the arguments that follow.

The enablement requirement is embodied in the first paragraph of 35 U.S.C. § 112, which requires the specification to describe the invention "in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same." 35 U.S.C. § 112 (2006). As explained in MPEP § 2164.01, "even though the statute does not use the term 'undue experimentation,' it has been interpreted to require that the claimed invention be enabled so that any person skilled in the art can make and use the invention without undue experimentation." MPEP § 2164.01 (citing *In re Wands*, 858 F.2d at 737, 8 U.S.P.Q.2d at 1404 (Fed. Cir. 1988)). "The test of enablement is not whether any experimentation is necessary, but whether, if experimentation is necessary, it is undue." MPEP § 2164.01 (citing *In re Angstadt*, 537 F.2d 498, 504, 190 U.S.P.Q. 214, 219 (CCPA 1976). That determination is made from the viewpoint of persons experienced in the field of the invention, *Elan Pharm., Inc. v. Mayo Found.*, 346 F.3d 1051, 68 U.S.P.Q.2d 1373 (Fed. Cir. 2003), and "requires the application of a standard of reasonableness, having due regard for the nature of the invention and the state of the art." *In re Wands*, 858 F.2d 731, 737, 8 U.S.P.Q.2d 1400 (Fed. Cir. 1988).

The undue experimentation inquiry invokes several factors, including, but not limited to: (1) the breadth of the claims; (2) the nature of the invention; (3) the state of the prior art; (4) the level of one of ordinary skill; (5) the level of predictability in the art; (6) the amount of direction provided by the inventor; (7) the existence of working examples; and (8) the quantity of experimentation needed to make or use the invention based on the content of the disclosure. *In re Wands,* 858 F.2d 731, 737, 8 U.S.P.Q.2d 1400 (Fed. Cir. 1988); MPEP § 2164.01(a). As

demonstrated below, applying these factors to the present case leads to the conclusion that the skilled artisan could have practiced the claimed methods without an undue amount of experimentation and, therefore, that the enablement rejection of claims 1, 4-10, 12-16, 18-19, and 50-51 is improper, and should be reversed.

1. Claims 1, 4-10, 12-16, 18-19, and 50 are enabled.

a. The breadth of the claims is commensurate with the specification.

Claims 1, 4-10, 12-16, 18-19, and 50 are directed to methods that comprise recombinant protein production in "non-human" host cells. The specification enables those skilled in the art to practice such methods without an undue amount of experimentation.

For example, the specification discloses numerous expression vectors containing constitutive or inducible promoters directing expression of the fusion protein in a particular host cell. *See* Specification, page 7, lines 8-37. In addition to describing the components of a non-plant expression vector, such as *E. coli* expression vectors pTrc and pET, the specification describes components of a typical plant expression vector, such as plant-specific promoters and selection markers, at page 8, lines 8-31.

The specification also illustrates several methods for introducing an expression vector into a host cell. For example, in the case of plant host cell, the specification discloses several gene transfer methods, including but not limited to *Agrobacterium*-mediated transformation, biolistics, electroporation, and PEG-mediated uptake. *See* Specification at page 9, lines 13-20. Furthermore, the application at page 9, lines 27-32, teaches methodology for regenerating a plant from a transformed cell and identifying transgenic plants expressing the recombinant protein of interest.

The specification discloses exemplary methods for recovering and purifying the recombinant protein of interest. For example, the specification states that a cell extract

containing an expressed pro-peptide-heterologous fusion protein can be applied to a chromatographic column. *See* Specification, page 14, lines 6-14. Selective binding of the fusion protein to antibodies raised against the pro-peptide sequence and immobilized onto the column results in selective retention of the fusion protein. *Id.* As an alternative to antibodies, the specification teaches that a gene encoding another immunogenic domain or a gene encoding a peptide with affinity for a commonly used column material, such as cellulose, glutathione-S-transferase or chitin, or any other desirable tag, may be included in the fusion protein. *Id.* at lines 10-15. The specification proffers yet another purification method, in which a peptide encoding a sequence which results in anchoring of the fusion protein in the cell wall is included in the fusion protein construct. Expression of such a fusion protein results in immobilization of the protein of interest to the cell wall, permitting its isolation by cleaving the anchored protein, washing the cells with water or washing buffer, centrifugation, and obtaining the protein from the washing buffer. *Id.* at lines 15-22.

The specification includes working examples illustrating the claimed methods. Example 1 reports the production of hirudin in *E. coli* cells (bacterial host cells) using chymosin as the autocatalytically maturing aspartic protease, while Example 2 reports a similar method for the production of carp growth hormone. Example 3 reports the production of carp growth hormone using a gut extract from red turnip beetle as the source of the autocatalytically maturing aspartic protease, to simulate *in vivo* cleavage conditions. *See* Specification, pages 14-17.

Thus, the specification provides ample disclosure to guide the skilled artisan in the recombinant production of a protein in non-human host cells.

b. The state of the art of recombinant protein production was advanced as of the priority date.

The Examiner questions the enablement of the invention with regard to recombinant protein production in various host cells, but recombinant protein production *per se* is not the focus of the invention. Instead, as discussed above, the invention relates to the discovery that a

recombinant polypeptide of interest can be obtained by methods comprising transforming a host cell with an expression vector comprising a chimeric nucleic acid sequence that encodes a fusion protein that comprises (i) a full-length chymosin pro-peptide and (ii) the recombinant polypeptide of interest, and contacting the fusion protein with a mature form of an autocatalytically maturing aspartic protease that is capable of cleaving the chymosin pro-peptide from the fusion protein to release the recombinant polypeptide of interest. *See* Specification, page 3, line 35-page 4, line 18. Indeed, the specification teaches at page 7, lines 6-7, that "[t]he type of host cell which is selected to express the fusion protein is not critical to the present invention and may be as desired." Thus, the enablement rejection is founded on an improper evaluation of the nature of the invention.

As explained below, at the time of the invention, the state of the art of the production of heterologous proteins in non-human hosts was advanced and predictable. The evidence of record demonstrates that the state of the art of recombinant protein production art was advanced at the time the application was filed. That is, as of the filing date of the instant invention, recombinant protein production was well known and practiced in a variety of hosts including but not limited to animals, bacteria, insects, and plants. *See, e.g.*, Dyck *et al., Trends in Biotechnology* 21:9:394-399 (2003). Numerous pre-filing publications of record teach recombinant protein production in a variety of non-human hosts. In view of this evidence, Appellants believe that the Examiner's concerns surrounding the enablement of the invention with regard to the production of the recited fusion protein in a non-human host cell are misplaced and erroneous.

(i) Animal host cells

Several scientific articles and patent documents of record evidence the advanced state of recombinant protein production in animals. For example, Aigner *et al.*, *Transgenic Research* 5: 405-411 (1996), demonstrates the expression of a tyrosinase in transgenic rabbits. Butler *et al., Thrombosis and Haemostasis* 78(1): 537-542 (1977), demonstrates the expression of human fibrinogen in the milk of transgenic animals. In a similar vein, U.S. Patent No. 5,959,171 (priority date December 20, 1993, issued October 27, 1998) discloses a method for producing

erythropoietin in the milk of a transgenic non-human mammal, with specific examples relating to the production of erythropoietin in mice. U.S. Patent No. 5,827,690 (priority date August 17, 1994, issued September 28, 1999) discloses a method for expressing heterologous and assembled immunoglobulins in the milk of a transgenic mammal, with specific examples relating to the production of immunoglobulins in mice.

Moreover, a 1992 review article by Jänne *et al.* summarizes the progress and feasibility of the production of therapeutic or industrial proteins in animals, demonstrating that the field of recombinant protein production in animals had rapidly advanced by the early 1990s. Jänne *et al.*, *Annals of Medicine* 24: 273-280 (1992). As evidenced in the article, researchers had successfully produced a wide array of recombinant proteins in animals, including human α1-antitrypsin, metallothionein growth hormone, and human antihaemophilic factor IX in sheep; human tissue plasminogen activator in goats; human lactoferrin and human erythropoietin in cattle, and human erythropoietin in mice.

Clearly, as of the invention's priority date, the state of recombinant protein production in animals was quite advanced.

(ii) Bacterial host cells

Just as recombinant protein expression in animals was routine in the art as of the instant application's priority date, so too was recombinant protein expression in bacteria. This is evidenced by pre-filing date publications of record, such as the review article by Sawers & Jarsch, *Appl Microbiol Biotechnol* 46: 1-9 (1996). As displayed in Table 1, at page 2 of the review article, a number of proteins have been expressed in bacterial systems, including proteases and lypases, glycose oxidase, cholesterol oxidase, glycose dehydrogenase, penicillin-G acylase and human therapeutics like recombinant tissue plasminogen activator and insulin. In addition to disclosing bacterially-expressed proteins, Sawers & Jarsch compare a number of bacterial expression systems and describe how they can be applied to the industrial-scale production of recombinant proteins.

Supplementing this already information-rich area of art, the specification itself teaches several methods for maximizing recombinant protein expression in *E. coli*. For example, the specification teaches that "[o]ne strategy to maximize recombinant protein expression in *E. coli* is to express the protein in host bacteria with an impaired capacity to proteolytically cleave the recombinantly expressed proteins." *See* Specification, page 7, lines 21-29. The specification teaches that "[a]nother strategy is to alter the nucleic acid sequence of the chimeric DNA to be inserted into an expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins." *See* Specification, page 7, lines 21-29.

Thus, the state of recombinant protein production in bacteria was advanced as of the invention's priority date, and the specification provides further guidance in this regard.

(iii) Insect host cell

Recombinant protein production in insects was advanced at the time of priority date of the instant application. Several scientific articles and patent documents of record evidence the advanced state of recombinant protein production in insects. For example, Duncker *et al.*, *Transgenic Research* 5: 49-55 (1996), reports expression of an antifreeze protein in *Drosophila*, and Yeh *et al.*, *PNAS* 92: 7036-7040 (1995) reports expression of the Green Fluorescent Protein in *Drosophila.* Similarly, U.S. Patent No. 5,472,858 (priority date June 4, 1991, issued December 5, 1995) discloses recombinant protein production in insect larvae.

Additionally, the instant specification discloses publicly accessible insect expression systems for producing recombinant proteins. "Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9 cells) include the pAc series (Smith *et al.*, (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow, V. A., and Summers, M. D., (1989) *Virology*)." *See* Specification, page 7, lines 34-36.

Thus, the evidence of record demonstrates that recombinant protein in an insect system was well within the capabilities of the skilled artisan as of the invention's priority date.

(iv) Plant host cells

Recombinant protein production in plants also was advanced at the time of filing the instant application. Several scientific articles and patent documents of record evidence the advanced state of recombinant protein production in plants. Nearly ten years before Appellants' present invention, researchers already had expressed proteins in plants. For example, Hiatt, *Nature* 344: 469-470 (1990), discloses expressing antibodies in plants. A review article, Mason & Arntzen, *Tibtech* 13: 388-392 (1995), discusses results from using plants as a vehicle to produce vaccines, including the expression of *Streptococcus mutans* spaA protein in tobacco, and the expression of *E. coli* heat-labile enterotixin B subunit and *E. coli* cholera-toxin B subunit in tobacco and potato. Likewise, Lyons *et al.*, *Pharmaceutical News* 3(3): 7-12 (1996), discusses producing protein pharmaceuticals in transgenic plants. Table 1 of Lyons discloses peptides or proteins that have been expressed in transgenic plants. Exemplary plant-produced proteins include 9 vaccines (such as hepatitis B surface antigen, Norwalk virus capsid protein, and human immunodeficiency virus); 5 antibodies (including mouse catalytic antibody 6D4 and Mouse Mab B 1-8; 2 serum proteins including human serum albumin and human protein C); 2 cytotoxins including α-trichosanthin and ricin; and the neuropeptide human epidermal growth factor. *Id* at Table 1, page 7.

In addition to the wealth of scientific literature teaching how to produce recombinant proteins in plants, several issued U.S. patents disclose methods for recombinant plant protein production. For example, U.S. Patent No. 5,650,554 (priority date February 22, 1991, issued July 22, 1997), issued to the assignee of the present invention, discloses a method for expressing recombinant polypeptides in a plant, such as tobacco and *Brassica napus*, or a bacterial host cell. Likewise, U.S. Patent No. 5,639,947 (priority date October 27, 1989, issued June 17, 1997) discloses a transgenic plant comprising immunoglobulins.

Thus, the evidence of record demonstrates that recombinant protein production in plants was well within the capabilities of the skilled artisan as of the invention's priority date.

c. The specification should not disclose what is well-known.

As set forth in MPEP § 2164.03, the "more that is known in the prior art about the nature of the invention . . . the less information needs to be explicitly stated in the specification." Indeed, MPEP § 2164.05(b) provides that "[t]he specification need not disclose what is well-known to those skilled in the art and preferably omits that which is well-known to those skilled and already available to the public." In keeping with these guidelines, the instant specification need not provide, and preferably omits, detailed teachings and examples of recombinant protein production in various host cells, because that information was known in the art and readily available to the skilled artisan at the time of filing, as demonstrated above.

d. The level of skill in the art was advanced as of the priority date.

The evidence discussed above demonstrating the advanced state of the art of recombinant protein production also demonstrates that the level of skill in the art as of the priority date was advanced. No evidence of record suggests otherwise.

e. The level of predictability in the art is supplemented by the amount of guidance provided in the specification.

The evidence discussed above demonstrating the advanced state of the art of recombinant protein production indicates that there is a high level of predictability in the art. Moreover, Vain *et al.* (2002), a reference cited by the Examiner, demonstrates in a monocot plant, which is generally held to be more difficult to transform than a dicot plant, that 100% of 95 independently transformed rice plants successfully expressed one of two transgenes, and 87% expressed both transgenes. *See* Vain, Abstract and page 880, first paragraph of "Results and Discussion." Thus, Vain demonstrates a high level of predictability in recombinant protein production, even in a particularly difficult plant system background,.

This general level of predictability in the art of recombinant protein production is supplemented by the guidance provided in the specification. For example, as of the filing date of

the instant application, methods of plant transformation were well known in the art, and the specification illustrates a variety of known methods and vectors for introducing a transgene into a plant cell. Thus, the specification discloses several gene transfer methods, including but not limited to *Agrobacterium*-mediated transformation, biolistics, electroporation, and PEG-mediated uptake. *See* Specification at page 9, lines 13-20. Furthermore, the specification at page 9, lines 27-32, teaches methodology for regenerating a plant from a transformed cell and identifying transgenic plants expressing the recombinant protein of interest. The specification also teaches optimization of a chymosin pro-peptide sequence specific for plant codon usage, to overcome difficulties that can be encountered when expressing a heterologous (non-plant) sequence in a plant. Thus, the specification discloses a nucleic acid sequence encoding a full-length chymosin pro-peptide sequence that is expressed efficiently in plant cells. See, *e.g.*, the sequence set forth in Figure 1.

Accordingly, the predictability of the recombinant protein production methods known in the art is supplemented by the guidance provided in the specification.

f. The specification includes working examples that further enable practice of the claimed methods.

The specification includes several working examples that further enable the claimed invention.

Example 1 reports the use of the claimed methods to produce hirudin, which is an anticoagulant that inhibits thrombin. The example describes the construction of an expression vector that comprised a nucleic acid encoding a calf chymosin pro-peptide and hirudin variant 1 fusion protein, denoted "GST-Pro-Hirudin." See Specification, page 14, lines 30-27. The expression vector was used to transform *E. coli* cells (bacterial host cells) as described on page 15, lines 3-14. As described on page 15, lines 15-28, hirudin was released from the fusion protein by treatment with the autocatalytically maturing aspartic protease chymosin at pH 2.0. Cleavage occurred at the Phe43-Val44 peptide bond, which separates the chymosin pro-peptide

from mature hirudin. Specification, page 16, lines 4-5. As reported in Table 1 of the example, the method successfully yielded hirudin and the recombinant hirudin protein demonstrated anti-thrombin activity.

Example 2 reports the use of the claimed methods to produce carp growth hormone. The example at page 16, lines 14-22, describes the construction of an expression vector that comprised a nucleic acid encoding a calf chymosin pro-peptide and carp growth hormone fusion protein, denoted as "His-Pro-cGH" fusion protein. The expression vector was used to transform *E. coli* cells (bacterial host cells). As described at pages 16-17 of the specification, the fusion protein was purified by chelating affinity chromatography using Hi-Trap metal binding columns that were saturated with Zn^{+2} ions to affinity purify the fusion protein. In order to cleave the fusion protein, thereby releasing the recombinantly produced cap growth hormone protein, the fusion protein was cleaved with the autocatalytically maturing aspartic protease chymosin in phosphate buffer (pH 2). *See* Specification, page 17, lines 5-15. As reported in Figure 3, the carp growth hormone obtained by this method had the expected molecular weight of carp growth hormone (approximately 22 kDa). *Id.*

Because the specification describes several methods for practicing the claimed invention, the enablement requirement is satisfied. *See, e.g., In re Fisher*, 427 F.2d 833, 839, 166 U.S.P.Q. 18, 24 (CCPA 1970) (noting that as long as the specification discloses at least one method for making and using the claimed invention, the enablement requirement is satisfied).

g. There is no evidence that an undue quantity of experimentation would be required to practice the claimed methods.

The factors discussed above support the enabling quality of the instant specification with respect to the claimed methods. The Examiner has not cited any evidence that an undue quantity of experimentation would be required to carry out the claimed methods.

While the Examiner has proffered three articles (Dyck, Vain, and Potrykus) in the Final Office Action of July 12, 2006, each allegedly discussing ongoing issues in the field of recombinant protein production, none of these articles suggest that an undue amount of experimentation is required for someone skilled in the art to produce a recombinant protein in a non-human host. Instead, these articles support Appellant's position as to the advanced state of the art, because they address issues that have arisen as recombinant protein production has moved out of the research laboratory into the commercial world.

The Examiner cites Dyck for stating that "the generation of transgenic domestic animals is difficult . . ." and that "current methods . . . are relatively inefficient and time-consuming" (emphasis added). *See* Dyck *et al.* (2003), at page 396, left column. These alleged problems are purely commercial considerations that have no bearing on the enablement of the claimed methods. Indeed, the entire focus of Dyck is the evaluation of various transgenic systems as "bioreactors" for large-scale production of proteins, and Dyck starts with the premise that "[t]he ability of transgenic animals to produce complex, biologically active recombinant proteins . . . has stimulated a great deal of interest in this area." *See* Dyck *et al.* (2003), Abstract.

Dyck itself cites numerous categories of successful transgenic protein production, including transgenic milk ("Foreign proteins are commonly reported to be expressed in transgenic milk at rates of several grams per litre," pg. 395); the blood of transgenic pigs ("[P]igs producing human haemoglobin in their own circulatory system have been produced," pg. 395); methods using retroviruses ("[R]etroviruses have been used to successfully produce transgenic mice and viral integration of recombinant sequences into bovine embryos to produce transgenic calves have been reported," pg. 396); methods using embryonic stem (ES) cells or primordial germ (PG) cells ("Reviews of the literature indicate that the production of chimeric animals with ES or PG cell technology has been applied successfully in mice, rabbits, pigs, cattle and poultry," pg. 397), and methods using pronuclear microinjection, which Dyck characterizes as being "the most straightforward and consistently successful means of gene transfer for most species" (pg. 397). This is hardly evidence of non-enablement.

Vain is cited for stating that "transgene expression in plants remains largely unpredictable," but that statement is taken out of context. *See* Vain *et al.* (2002), at page 878, first paragraph of "Introduction." Reading the complete sentence (at pg. 878, col. 2) reveals that the "unpredictable" factors being noted are "variation in expression levels and stability between independently transformed plants," not the ability to achieve transgenic expression *per se*. In Vain's own experiments, 100% of 95 independently transformed rice plants successfully expressed one of two transgenes, and 87% expressed both transgenes. *See* Vain, Abstract. Thus, Vain does not support an assertion that undue experimentation is required to produce transgenic proteins in plants.

The Examiner cites Potrykus for the premise that gene transfer in cereals is largely unsuccessful, but his reliance on this article to support an enablement rejections is misplaced. Potrykus presents an admittedly "subjective" review of different methods that have been used to effect gene transfer in cereal crops "and their potential agronomic utility" (emphasis added). *See* Potrykus (1990) at page 535, top of left column. Thus, this article, like those discussed above, relates to issues encountered on the road to commercialization.

While Potrykus criticizes a number of different transgenic methods, it also acknowledges methods that have proven successful, including methods using *Agrobacterium* or agroinfection to transform dicots and methods using protoplasts for direct gene transfer of cereals. *See* Potrykus, page 538, right column entitled "A routine and efficient method for the production of transgenic plants from numerous non-cereal species." In fact, the "Note added in proof" at page 542 reports Potrykus' own work to establish "what we believe is proof of the recovery of transgenic offspring of *Indica*-type rice." That work, reported in Datta *et al., Bio/Technology* (1990) 736-40, is said to result in "a simple and reproducible method of transformation of an important food crop." Thus, Potrykus does not undermine the enablement of the present invention with respect to recombinant protein production in plants.

Appellants also question the validity of Potrykus as an accurate reflection of the state of the art. As noted above, Potrykus itself states it presents a "subjective" discussion. Moreover,

Potrykus admits that it includes "several statements . . . for which no solid experimental data are available." *See* Potrykus, page 535, right column. The evidentiary value of Potrykus' critiques therefore pales in view of the numerous references of record that present scientific data demonstrating the successful production of transgenic proteins in plants, including Potrykus' own work.

In citing Dyck, Vain and Potrykus, the Examiner loses sight of the fact that § 112 does not require an applicant "to enable one of ordinary skill in the art to make and use a perfected, commercially viable embodiment." MPEP § 2164 (citing *CFMT, Inc. v. Yieldup Int'l Corp.*, 349 F.3d 1333, 1338, 68 U.S.P.Q.2d 1940, 1944 (Fed. Cir. 2003). All that § 112 requires is that the specification enable those skilled in the art to practice the claimed invention without an undue amount of experimentation. The fact that some experimentation may be necessary to practice the invention does not render the specification non-enabling. "[A] considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed." *In re Wands*, 858 F.2d 731, 737, 8 U.S.P.Q.2d 1400, 1404 (Fed Cir. 1988). Because Applicant has demonstrated that recombinant protein production in a wide variety of non-human hosts (including animals, bacteria, insects and plants) was well-developed at the time the present application was filed, the instant enablement rejection is improper, and should be reversed.

2. Claim 51 is enabled.

Claim 51 is largely identical to claim 1, but recites embodiments where the host cell is a bacterial cell, yeast cell or plant cell. The arguments set forth in Section B.1 above with reference to claims 1, 4-10, 12-16, 18-19, and 50 support the enablement of claim 51. Appellants separately argue the patentability of claim 51, however, because the Examiner has admitted enablement with respect to recombinant protein production in bacterial and yeast cells, and because the record directly supports enablement with respect to plant cells.

a. The Examiner has admitted that claim 51 is enabled with respect to recombinant protein production in bacteria and yeast.

In the Office Action mailed July 12, 2006, page 6, the Examiner recognizes that recombinant production in bacteria, insects and yeast is enabled:

> "There is no dispute that the fusion protein set forth in the claims could be recombinantly produced in a bacteria, yeast, or insect cell at the time of the invention without requiring undue experimentation."

This admission, coupled with the record evidence of enablement discussed in section B.1 above, demonstrates that claim 51 is enabled with respect to recombinant protein production in bacterial and yeast cells.

b. The record evidence demonstrates that claim 51 is enabled with respect to recombinant production in plants.

Section B.1 above demonstrates that the invention is enabled with respect to recombinant protein production in non-human host cells, including plants. The record evidence discussed in Section B.1.b.iv specifically supports the enablement of claim 51 with respect to recombinant protein production in plants. This evidence includes Hiatt, *Nature* 344: 469-470 (1990) (expression of antibodies in plants); Mason & Arntzen, *Tibtech* 13: 388-392 (1995) (production of vaccine antigens in plants); Lyons *et al.*, *Pharmaceutical News* 3(3): 7-12 (1996) (production of protein pharmaceuticals in plants); U.S. Patent No. 5,650,554 (expression of recombinant polypeptides in plants); U.S. Patent No. 5,639,947 (production of immunoglobulins in plants). Moreover, Section B.1.e. illustrates teachings in the specification that provide further guidance on recombinant protein production in plants, such as several gene transfer methods (*see* Specification at page 9, lines 13-20), methodology for regenerating a plant from a transformed cell and identifying transgenic plants expressing the recombinant protein of interest (*see* Specification at page 9, lines 27-32), optimization of a chymosin pro-peptide sequence specific for plant codon usage (*see, e.g.*, the sequence set forth in Figure 1).

Taken as a whole, the evidence of record demonstrates that recombinant protein production in plants was well within the capabilities of the skilled artisan as of the invention's priority date. Thus, the enablement rejection of claim 51 is improper, and should be reversed.

C. The instant claims are not obvious in view of the cited references.

The requirement for non-obviousness is embodied in 35 U.S.C. § 103, which provides in pertinent part that:

> A patent may not be obtained . . . if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill....

Obviousness is a question of law which is based upon several factual inquires related to the scope and content of the prior art, the differences between the prior art and the claims, and the level of ordinary skill in the art. *See*, e.g. *In re Kotzab*, 217 F.3d 1365, 1375 (Fed. Cir. 2000). Obviousness may be based on the teachings of a single reference or the combined teachings of several references.

A legally sufficient or "*prima facie*" case of obviousness must satisfy three basic criteria. First, the prior art reference, or combination of references, must teach or suggest each of the claim limitations. *In re Royka*, 490 F.2d 981, 985, 180 U.S.P.Q. 580 (CCPA 1974). Second, the prior art, in light of knowledge generally available to one of ordinary skill in the art, must evidence some suggestion or motivation to modify the reference or to combine reference teachings to arrive at the claimed invention. *In re Vaeck*, 947 F.2d 488, 493, 20 U.S.P.Q.2d 1438 (Fed. Cir. 1991). Third, the prior art must evidence that those skilled in the art would have had a reasonable expectation of success in achieving the invention. *Id.* at 493.

Simply because references can be combined to produce the claimed invention is not sufficient to establish obviousness. *In re Mills*, 916 F.2d 680 (Fed. Cir. 1990). "Both the suggestion and the reasonable expectation of success must be founded in the prior art, not the applicant's disclosure," *Vaeck*, 947 F.2d at 493, in order to avoid the "insidious effect of a

hindsight syndrome wherein that which only the invention taught is against its teacher," *Kotzab*, 217 F.3d at 1369.

Further, the reasonable expectation of success must be judged from the time the invention was made, and from the perspective of someone of ordinary skill in the art. *Micro Chem., Inc. v. Great Plains Chem Co.*, 103 F.3d 1538, 1547, 65 USLW 2487, 41 U.S.P.Q.2d 1238 (Fed. Cir. 1997). "That the inventors were ultimately successful is irrelevant to whether one of ordinary skill in the art at the time the invention was made would have reasonably expected success." *Life Tech., Inc. v. Clontech Labs., Inc.*, 224 F.3d 1320, 1326 (Fed. Cir. 2000). Thus, a reasonable expectation of success must be (a) grounded in the prior art; (b) determined at the time of invention; and (c) assessed from the vantage point of one of ordinary skill in the art.

The references cited by the Examiner fail to teach or suggest a method of preparing a recombinant polypeptide of interest that comprises transforming a host cell with an expression vector comprising a chimeric nucleic acid sequence that encodes a fusion protein that comprises (i) a full-length chymosin pro-peptide and (ii) the recombinant polypeptide of interest, and contacting the fusion protein with a mature form of an autocatalytically maturing aspartic protease that is capable of cleaving the chymosin pro-peptide from the fusion protein, to release the recombinant polypeptide of interest. There is no motivation in the cited references, or any other evidence of record, to combine the references in the manner asserted in order to arrive at the claimed invention, and no combination of cited references provides any reasonable expectation that such a method would result in production of recombinant polypeptide of interest.

1. Claims 1, 4, 6-9, 13, 15, 19 and 51 are not obvious over Ward in view of Walsh and Yonezawa.

The rejected claims are directed to methods of preparing a recombinant polypeptide of interest that comprises transforming a host cell with an expression vector comprising a chimeric nucleic acid sequence that encodes a fusion protein that comprising (i) a full-length chymosin pro-peptide and (ii) the recombinant polypeptide of interest, and contacting the fusion protein

with a mature form of an autocatalytically maturing aspartic protease that is capable of cleaving the chymosin pro-peptide from the fusion protein to release the recombinant polypeptide of interest. Such methods are not taught or suggested by Ward in view of Walsh and Yonezawa.

Ward describes a nucleic acid encoding a fusion protein that includes a bovine chymosin prosequence as a cleavable linker. The reference provides no specific teachings of how to cleave the bovine chymosin prosequence from its fusion protein, and certainly does not teach or suggest using an autocatalytically maturing aspartic protease to do so. Walsh discloses the use of a bovine k-casein sequence that is sensitive to chymosin cleavage as a cleavable linker for fusion proteins, but does not teach or suggest the use of a chymosin pro-peptide for such a purpose. Yonezawa reports the various cleavage sites for several chromogenic chymosin substrates. This combination of references does not suggest the use of a mature form of an autocatalytically maturing aspartic protease to cleave a chymosin pro-peptide from a fusion protein comprised of a heterologous protein.

In rejecting the claims, the Examiner states in the Final Office Action of July 12, 2006, in the paragraph bridging pages 10 and 11, that:

> "C-terminal amino acid of a chymosin pro-peptide is a Phe, and one of ordinary skill in the art would recognize that Met is usually the first amino acid of a given polypeptide. Although the claims are not so limited, the Examiner has directed the rejection to the recited fusion protein having a Phe-Met junction between the chymosin pro-peptide and the heterologous protein."

The Examiner concludes on page 12:

> "One would have been motivated to use chymosin as the fusion protein-cleaving agent in the method of Ward *et al.* because Walsh *et al.* teaches that a Phe-Met site is the specific cleavage site of k-casein and both Walsh *et al.* and Yonezawa *et al.* demonstrate that chymosin can cleave a Phe-Met site. One would have a reasonable expectation of success for practicing the method for fusion protein preparation and cleavage of Ward *et al.* using mature chymosin as

> the fusion protein cleaving agent because of the results of Ward *et al.*, Walsh *et al.*, and Yonezawa *et al.*"

Thus, this rejection is based on the Examiner's incorrect assumption that chymosin will cleave any fusion protein at a Phe-Met bond. As explained below, chymosin does not cleave any and all fusion proteins at a Phe-Met bond. The error of the Examiner's assumption is shown by the specification and the previously submitted Declaration under 37 C.F.R. § 1.132 of Dr. Maurice Moloney, a named inventor of the application, and the scientific references discussed therein.

In his Declaration, Dr. Moloney acknowledges that prior art references show that chymosin can cleave the substrate κ-casein at a specific Phe-Met bond, but he explains that the references also teach that the primary structure of the amino acids surrounding the Phe-Met bond is essential to the cleavage reaction. Specifically, the art teaches that a minimum chain length of five amino acid residues including the sequence Ser-Phe-Met-Ala is essential to bring about a cleavage of the Phe-Met bond in κ-casein. (See paragraph 6 of the Moloney Declaration, citing Visser *et al.* and Schattenkerk *et al.*). In addition, in paragraph 8 of his Declaration, Dr. Moloney refers to Example 1 and Figure 1 of the application as filed, and explains how they demonstrate that the Phe-Met bond that is present in the GST-Pro-Hirudin fusion protein is not cleaved by chymosin. Instead, cleavage occurs between a Phe-Val bond in that fusion protein. Dr. Moloney explains with reference to Figure 2 that cleavage of the His-Pro-cGH fusion protein also does not occur at a Phe-Met bond, but rather at a Phe-Ser bond. Thus, the Moloney declaration and the prior art and application data discussed therein demonstrate that chymosin does not cleave proteins at any and all Phe-Met bonds.

Thus, while the Examiner alleges that the invention is obvious because it would have been expected that chymosin will cleave the Phe-Met bond between a chymosin pro-peptide and heterologous peptide, the evidence of record (including the specification, Moloney Declaration, and scientific references) clearly demonstrates that this is not the case, *e.g.*, it was known in the art that chymosin does not cleave any and all Phe-Met junctions. Moreover, Figures 1-2 of the

specification reveal that, for these fusion proteins, cleavage does not occur at Phe-Met bonds. Thus, the references cited by the Examiner do no suggest the present invention, or provide any expectation of being able to obtain a recombinant polypeptide of interest by the recited methods.

There is simply no teaching or motivation in the cited art of preparing a recombinant polypeptide of interest by producing a fusion protein comprising a chymosin pro-peptide and the polypeptide of interest, and using a mature aspartic protease to cleave the chymosin pro-peptide sequence from the fusion protein to release the recombinant polypeptide of interest, as claimed. One of ordinary skill in the art would not have had a reasonable basis for expecting that an aspartic protease would be capable of cleaving a chymosin pro-peptide from a fusion protein to release the recombinant polypeptide, and did not know, for example, whether the aspartic protease would cleave the recombinant polypeptide at undesired sites and/or would cleave off too many or too few amino acid residues around the junction between the pro-peptide and the recombinant polypeptide. Without an assurance of accurate cleavage, there was no motivation to have employed an aspartic protease as presently claimed.

Because it is only the instant specification that recognizes and teaches that aspartic proteases are capable of cleaving a chymosin pro-peptide from a fusion protein to release a recombinant polypeptide of interest, these obviousness rejections are improperly founded and should be reversed.

2. Claim 5 is not obvious over Ward in view of Walsh and Yonezawa further in view of Fine.

Claim 5 depends from claim 1, and recites that the recombinant polypeptide is hirudin or carp growth hormone. Such a method is not taught or suggested by Ward in view of Walsh, Yonezawa, and Fine.

For reasons already advanced, none of Ward, Walsh, and Yonezawa, alone or together, teach or suggest the method recited in claim 1, from which claim 5 depends. Fine is cited for allegedly disclosing recombinant expression of carp growth hormone. Fine's teachings,

however, do not remedy the inability of Ward, Walsh, and Yonezawa to render obvious the claimed invention. For example, Fine does not teach or suggest using a mature form of an autocatalytically maturing aspartic protease to cleave a chymosin pro-peptide from a fusion protein comprising carp growth hormone, or provide any expectation of success in being able to do so. Since Fine does not remedy the deficiencies of Ward, Walsh, and Yonezawa, this combination of references does not render the present invention obvious. As there is no *prima facie* obviousness, the rejection is improper and should be reversed.

3. Claims 14 and 50 are not obvious over Ward in view of Walsh and Yonezawa further in view of Dunn.

Claims 14 and 50 depend from claim 1, and recite specific embodiments of the aspartic protease of step d). Claim 14 recites that the aspartic protease is heterologous to the chymosin pro-peptide and claim 50 recites that the aspartic protease is pepsin. Such methods are not taught or suggested by Ward in view of Walsh, Yonezawa, and Dunn.

For reasons already advanced, none of Ward, Walsh, and Yonezawa, alone or together, teach or suggest methods of preparing a recombinant polypeptide of interest that comprises transforming a host cell with an expression vector comprising a chimeric nucleic acid sequence that encodes a fusion protein that comprising (i) a full-length chymosin pro-peptide and (ii) the recombinant polypeptide of interest, and contacting the fusion protein with a mature form of an autocatalytically maturing aspartic protease that is capable of cleaving the chymosin pro-peptide from the fusion protein to release the recombinant polypeptide of interest.

Dunn is cited for teaching that a number of aspartic proteases have the ability to proteolytically cleave a recognition site having Phe in the P1 position. However, the teachings of Dunn relating to the ability of mature aspartic proteases to cleave specific peptides at specific sites in no way teaches or suggests the claimed invention, which recites a method wherein a mature aspartic protease other than chymosin is contacted with a fusion protein comprising a

chymosin pro-peptide sequence and cleaves the chymosin pro-peptide from the fusion protein to release a recombinant polypeptide of interest.

As stated above, there simply is no hint in the prior art of using a mature aspartic protease to cleave a chymosin pro-peptide sequence from a fusion protein to release a recombinant polypeptide of interest. The fact that mature aspartic proteases have been shown to cleave specific peptides at specific sites in no way implicates the use of a mature aspartic protease in accordance with the present invention. As noted above, those skilled in the art had no reasonable basis for expecting that an aspartic protease would be capable of cleaving a chymosin pro-peptide from a fusion protein to release the recombinant polypeptide, and did not know, for example, whether the aspartic protease would cleave the recombinant polypeptide at undesired sites and/or would cleave off too many or too few amino acid residues around the junction between the pro-peptide and the recombinant polypeptide. Without an assurance of accurate cleavage, there was no motivation to have employed an aspartic protease as presently claimed, and no expectation of success in being able to do so.

Because it is only the instant specification that recognizes and teaches that aspartic proteases are capable of cleaving a chymosin pro-peptide from a fusion protein to release a recombinant polypeptide of interest, this obviousness rejection is improperly founded and should be reversed.

VIII. CONCLUSION

For the reasons discussed above, Appellants respectfully submit that all pending claims are in condition for allowance, and respectfully request that the rejections be reversed in whole, and that the claims be allowed to issue.

Respectfully submitted,

April 18, 2007
Date

for Micheline Gravelle
Registration No. 40,261

Reg No 37,288

Bereskin & Parr
Box 401, 40 King Street West
Toronto, Ontario
Canada, M5H 3Y2
Tel: 416-957-1682
Fax: 416-361-1398

Should additional fees be necessary in connection with the filing of this paper, or if a petition for extension of time is required for timely acceptance of same, the Commissioner is hereby authorized to charge Deposit Account No. 19-0741 for any such fees; and patentee hereby petition for any needed extension of time.

IX. APPENDIX A: CLAIMS APPENDIX

1. A method for the preparation of a recombinant polypeptide comprising

a) transforming a non-human host cell with an expression vector comprising:

(1) a nucleic acid sequence capable of regulating transcription in a host cell, operatively linked to

(2) a chimeric nucleic acid sequence that encodes a fusion protein, wherein said chimeric nucleic acid sequence comprises (a) a nucleic acid sequence encoding a full-length chymosin pro-peptide, linked in reading frame to (b) a nucleic acid sequence that is heterologous to the pro-peptide and that encodes the recombinant polypeptide, wherein the heterologous nucleic acid sequence is located immediately downstream of the nucleic acid sequence encoding the chymosin pro-peptide; operatively linked to

(3) a nucleic acid sequence encoding a termination region that is functional in said host cell,

b) growing the non-human host cell to produce said fusion protein,

c) obtaining said fusion protein from said non-human host cell, and

d) contacting said fusion protein with a mature form of an autocatalytically maturing aspartic protease that is capable of cleaving the chymosin pro-peptide, whereby said chymosin pro-peptide is cleaved from said fusion protein to release said recombinant polypeptide.

4. The method according to claim 1 wherein said aspartic protease of step d) is selected from the group consisting of chymosin, pepsin, pepsinogen, cathepsin and yeast proteinase A.

5. The method according to claim 1 wherein the recombinant polypeptide is hirudin or carp growth hormone.

6. The method according to claim 1 wherein the chimeric nucleic acid sequence does not include a sequence encoding a mature form of chymosin.

7. The method according to claim 1 wherein step d) is effected at a pH of from about 2 to about 7.

8. The method according to claim 7 wherein the pH is from about 2 to about 4.5.

9. The method according to claim 1 wherein step d) is effected in vitro.

10. The method according to claim 1 wherein step d) is effected in vivo.

12. The method according to claim 10 wherein step d) is effected in the milk, the stomach, or the gut of an animal.

13. The method according to claim 1 wherein the aspartic protease of step d) is chymosin.

14. The method according to claim 1 wherein the aspartic protease of step d) is heterologous to the chymosin pro-peptide.

15. The method according to claim 13 wherein step d) is effected in vitro.

16. The method according to claim 13 wherein step d) is effected in vivo.

18. The method according to claim 16 wherein step d) is effected in the stomach, gut, or milk of an animal.

19. The method according to claim 1 wherein said nucleic acid sequences are deoxyribonucleic acid (DNA) sequences.

50. The method according to claim 1 wherein said aspartic protease of step d) is pepsin.

51. A method for the preparation of a recombinant polypeptide, comprising

a) transforming a host cell with an expression vector comprising:

(1) a nucleic acid sequence capable of regulating transcription in a host cell, operatively linked to

(2) a chimeric nucleic acid sequence that encodes a fusion protein, wherein said chimeric nucleic acid sequence comprises (a) a nucleic acid sequence encoding a full length chymosin pro-peptide, linked in reading frame to (b) a nucleic acid sequence that is heterologous to the pro-peptide and that encodes the recombinant polypeptide, wherein the heterologous nucleic acid sequence is located immediately downstream of the nucleic acid sequence encoding the chymosin pro-peptide; operatively linked to

(3) a nucleic acid sequence encoding a termination region that is functional in said host cell,

wherein the host cell is selected from the group consisting of bacterial cells, yeast cells and plant cells,

b) growing the host cell to produce said fusion protein;

c) contacting said fusion protein in vivo with a mature form of an autocatalytically maturing aspartic protease that cleaves the pro-peptide by expressing said autocatalytically maturing aspartic protease in said host cell,

whereby said pro-peptide is cleaved from said fusion protein to release said recombinant polypeptide.

X. APPENDIX B: EVIDENCE APPENDIX

Journal Articles

Aigner *et al.*, Expression of the Murine Wild-Type Tyrosinase Gene in Transgenic Rabbits. *Transgenic Research* 5: 405-411 (1996).
Entered by Appellants with Amendment filed April 21, 2006.

Butler *et al.*, Current Progress in the Production of Recombinant Human Fibrinogen in the Milk of Transgenic Animals. *Thrombosis and Haemostasis* 78(1): 537-542 (1997).
Entered by Appellants with Amendment filed April 21, 2006.

Datta *et al.*, Genetically Engineered Fertile Indicia-Rice Recovered From Protoplasts. *Bio/Technology* 736-40 (1990).
Entered by Appellants with Amendment filed September 12, 2006.

Duncker *et al.*, Expression Of A Cystine-Rich Fish Antifreeze In Transgenic *Drosophila melanogaster*. *Transgenic Research* 5: 49-55 (1996).
Entered by Appellants with Amendment filed April 21, 2006.

Dunn *et al.*, "Aspartic Proteinases," Advances in Experimental Medicine and Biology, Volume 362, Plenum Press, NY, 1995, pp. 1-9.
Entered by Examiner in Office Action mailed April 29, 2004.

Dyck *et al.*, Making Recombinant Proteins In Animals – Different Systems, Different Applications. *Trends in Biotechnology* 21:9:394-399 (2003).
Entered by Examiner in Office Action mailed July 12, 2006.

Fine *et al.*, Recombinant Carp (*Cyprinus Carpio*) Growth Hormone: Expression, Purification, And Determination Of Biological Activity. *General Comp. Endocrinol.* 89:51-61 (1993).
Entered by Examiner in Office Action mailed December 4, 2001.

Hiatt, A., Antibodies Produced In Plants. *Nature* 344: 469-470 (1990).
Entered by Appellants with Amendment filed April 21, 2006.

Jänne *et al.*, Transgenic Animals as Bioproducers of Therapeutic Proteins. *Annals of Medicine* 24: 273-280 (1992).
Entered by Appellants with Amendment filed April 21, 2006.

Lyons *et al.*, Production Of Protein Pharmaceuticals In Transgenic Plants. *Pharmaceutical News* 3(3): 7-12 (1996).
Entered by Appellants with Amendment filed April 21, 2006.

Mason H.S. and Arntzen C.J. Transgenic Plants As Vaccine Production Systems. *Tibtech* 13: 388-392 (1995).
Entered by Appellants with Amendment filed April 21, 2006.

Potrykus, I., Gene Transfer to Cereals: An Assessment. *Biotechnology* 8(6)535-542 (1990).
Entered by Examiner in Office Action mailed July 12, 2006.

Sawers, G and Jarsch, M., Alternative Regulation Principles For The Production Of Recombinant Proteins in *Escherichia coli*. *Appl Microbiol Biotechnol* 46: 1-9 (1996).
Entered by Appellants with Amendment filed April 21, 2006.

Vain, P *et al.*, Transgene Behavior Across Two Generations in a Large Random Population of Trasgenic Rice Plants Produced By Particle Bombardment. *Theor Appl Genet* 105:878-889 (2002).
Entered by Examiner in Office Action mailed f July 12, 2006.

Walsh *et al.*, Investigating the Use of The Chymosin-Sensitive Sequence of k-Casein as a Cleavable Linker Site in Fusion Proteins. *J. Biotech.* 45:235-241 (1996).
Entered by Examiner in Office Action mailed April 29, 2004.

Yeh *et al.*, Green Fluorescent Protein As A Vital Marker And Reporter Of Gene Expression In *Drosophila. PNAS* 92: 7036-7040 (1995).
Entered by Appellants with Amendment filed April 21, 2006.

Yonezawa *et al.*, Sensitive Fluormetric Assay for the Activity of Chymosin, *Int J Pept Protein Res* 47:56-61 (1996).
Entered by Examiner in Office Action mailed March 9, 2005.

Patent Documents

U.S. Patent No. 5,472,858
Entered by Appellants with Amendment filed April 21, 2006.

U.S. Patent No. 5,639,947
Entered by Appellants with Amendment filed April 21, 2006.

U.S. Patent No. 5,650,554
Entered by Appellants with Amendment filed April 21, 2006.

U.S. Patent No. 5,827,690
Entered by Appellants with Amendment filed April 21, 2006.

U.S. Patent No. 5,959,171

Entered by Appellants with Amendment filed April 21, 2006.

U.S. Patent No. 6,265,204

Entered by Examiner in Office Action mailed April 29, 2004.

Other Evidence

Rule 132 Declaration of Dr. Moloney (executed April 7, 2006)

Entered by Appellants with Amendment filed April 21, 2006; entered by Examiner as noted in Office Action mailed July 12, 2006.

XI. APPENDIX C: RELATED PROCEEDINGS APPENDIX

No related proceedings are pending.

Transgenic Research 5, 405–411 (1996)

Expression of the murine wild-type tyrosinase gene in transgenic rabbits

BERNHARD AIGNER[1]‡, URBAN BESENFELDER[1]‡, JANOS SEREGI[2], LASZLO V. FRENYO[2], TIBOR SAHIN-TOTH[2] and GOTTFRIED BREM[1,2]*

[1]*Institut für Tierzucht und Genetik, Veterinärmedizinische Universität Wien, Josef-Baumann-Gasse 1, A-1210 Wien, Austria (Fax: +43-1-25077 5692)*
[2]*Research Institute of the University for Veterinary Science, Üllő, Hungary*

Received 21 July 1995; revised 27 October 1995; accepted 6 November 1995

The tyrosinase gene is known to be essential for melanization and has been shown to rescue pigmentation in albino mice. Previously we have described the strict copy-number-dependent expression of a murine wild-type tyrosinase gene construct over several generations in transgenic mice. In this study, we analysed the same gene construct as a marker gene for the transmission and expression of transgenes in rabbits. Using an albino hybrid strain, we produced transgenic rabbits expressing the murine tyrosinase gene. Strict correlation between integration and expression of the transgene and stable germline transmission of the integrated gene construct according to the Mendelian pattern of inheritance was observed. Thus, breeding control was facilitated by simple phenotypic examination of the transgenic animals. In contrast to mice transgenic for the same gene construct, tyrosinase-transgenic rabbits showed a greater variety in hue, intensity and extent of coat pigmentation, which is caused by the diversity in the loci affecting the melanization. Benefits and limitations of tyrosinase as a marker gene for the detection of homozygous individuals in the albino hybrid strain used are discussed.

Keywords: breeding control; colour marker; gene transfer; pigmentation

Introduction

Pigmentation in mammals is based on the production of melanin, a heteropolymer of different metabolic intermediates of tyrosine. In mice, it is controlled by more than 50 independent loci consisting of more than 150 different alleles (Hearing, 1987). Tyrosinase (EC 1.14.18.1) is the key enzyme in this pathway and is exclusively expressed in melanocytes of the skin and pigment cell layers of the eyes (Silvers, 1979). In mice, the single copy gene maps to chromosome 7 and covers the *c*- (albino-) locus (Hearing and Tsukamoto, 1991). The murine tyrosinase gene has a length of about 70 kb and includes five exons coding for a mRNA of about 2 kb (Ruppert *et al.*, 1988). While the dominant wild-type allele *C* results in full tyrosinase activity, alleles giving rise to reduced melanin production have been described (Halaban *et al.*, 1988; Kwon *et al.*, 1989; Beermann *et al.*, 1990). The lack of enzyme activity is caused by the recessive allele *c* harbouring a defined point mutation in a highly conserved region of the first exon (Shibahara *et al.*, 1990). This point mutation was observed in all albino mouse strains examined (Jackson and Bennent, 1990). The lack of tyrosinase activity masks the information of all other loci affecting the melanization (Hogan *et al.*, 1986). After tyrosinase gene expression in tissue culture systems (Müller *et al.*, 1988; Takeda *et al.*, 1989; Yamamoto *et al.*, 1989) murine wild-type tyrosinase gene constructs have been shown to rescue the albino phenotype in mice (Beermann *et al.*, 1990; Tanaka *et al.*, 1990; Yokoyama *et al.*, 1990).

The molecular genetics of the albino phenotype in rabbits remain to be established. Nevertheless we have demonstrated the rescue of the albino phenotype of ZIKA^R hybrid rabbits by introducing a murine wild-type tyrosinase gene construct (Aigner and Brem, 1993). Having established the possibility for the use of the

‡Present address: Abteilung Biotechnologie in der Tierproduktion, IFA Tulln, Konrad-Lorenz-Str. 20, A-3430 Tulln, Austria.
*To whom correspondence should be addressed.

tyrosinase as a marker gene for the transmission of transgenes in co-integration experiments by strict copy-number dependent gene expression in mice (Aigner and Brem, 1994), in this study we test the same murine tyrosinase minigene construct (Beermann *et al.*, 1990) in another species for its use as a marker gene by analysing transmission and expression of the transgene in several transgenic rabbit lines.

Materials and methods

Generation of transgenic rabbits

The albino ZIKA^R hybrid strain was used for the generation of transgenic rabbits. Animal husbandry, embryo production, pronuclear microinjection of 500 copies of the transgene and laparoscopic transfer of the zygotes to the recipients were performed as previously described (Besenfelder and Brem, 1993; Brem, 1993). The gene construct ptrTyr4 has a length of 15.5 kb and includes 5.5 kb of the 5′ untranslated region, exon 1, almost the entire first intron and the exons 2–5 of the murine wild-type tyrosinase gene followed by the SV40 splice and polyadenylation cassette (Beermann *et al.*, 1990).

DNA analysis

Genomic DNA was isolated from tissue probes by standard protocols (Ausubel *et al.*, 1987). Polymerase chain reaction was carried out for the detection of transgenesis (Beermann *et al.*, 1990); transgenic animals showed a specific 515 bp signal. Examination of integration and stable transmission of the transgene by Southern analysis was performed as described (Aigner and Brem, 1994). A 1 kb SV40 fragment (*Bam* HI–*Sal* I) was used as a transgene specific probe (Gorman *et al.*, 1982).

The copy numbers of the transgenic rabbit lines were estimated by slot blot analysis. Dilutions of transgenic DNA (8, 4, 2 and 1 μg genomic DNA) were visually compared to differing amounts of the gene construct (in 4 μg rabbit DNA) representing various copy numbers and ptrTyr4-transgenic mouse probes with defined copy number (Aigner and Brem, 1994, 1995). Genomic DNA of non-transgenic animals served as negative control. The differentiation between homozygous and hemizygous littermates after mating two hemizygous siblings was done in duplicate by semiquantitative PCR as described (Aigner and Brem, 1995) using two different pairs of transgene-specific primers (Beermann *et al.*, 1990). The amount of DNA used in the semiquantitative PCR was evaluated by amplification of a 500 bp fragment deriving from the endogenous rabbit transferrin gene (Banfield *et al.*, 1991) with the primers RATF1 (5′-GCCTTTGTC-AAGCAAGAGACC-3′) and RATF2 (5′-CACAGCAGCT-CATACTGATCC-3′) at 62 °C annealing temperature.

RNA analysis

Total RNA was isolated from tissue probes by standard techniques (Chomcynski and Sacchi, 1987), mRNA by using the 'Quickprep Micro Purification' kit (Pharmacia). Reverse transcriptase (RT) PCR was performed as previously described (Innis *et al.*, 1990; Aigner and Brem, 1994). A 449 bp signal shows the specific expression of the integrated gene construct ptrTyr4.

Results and discussion

Generation and analysis of the transgenic founder rabbits

The albino ZIKA^R hybrid strain was used for the generation of rabbits transgenic for the murine wild-type tyrosinase gene construct ptrTyr4. On average, 15 micro-injected zygotes were transferred per uterus horn, resulting in 3.8 offspring born per recipient. The integration of the gene construct was examined by PCR. Table 1 shows the results of the tyrosinase gene transfer. A total of 27 animals were found having integrated the gene construct. The integration efficiency (transgenics/F_0 examined) was 11.4%, the overall efficiency (transgenics/zygotes transferred; corrected to the number of F_0 animals examined) was 0.66%. Comparable gene transfer efficiencies in rabbits were achieved in other programmes (Brem *et al.*, 1994). All phenotypically examined transgenic rabbit founders expressed the murine gene construct in coat and/or eyes, whereas none of the unpigmented animals harboured detectible parts of the gene construct. In mice, a high percentage of transgene expression was also achieved with different tyrosinase gene constructs (Yokoyama *et al.*, 1990; Overbeek *et al.*, 1991). Thus, we have rescued the albino phenotype in the ZIKA^R hybrid strain by tyrosinase gene transfer.

Table 1. Production of ptrTyr4-transgenic rabbits

Zygotes		*Recipients*		*Newborns*		*Transgenic by PCR*
microinjected	*transferred*	*total*	*pregnant*	*total*	*examined*[a]	
7890	7100 (90%)	229	107 (47%)	411	237	27 (11.4%)

[a]As consequence of health problems and death of newborns independent from the gene transfer programme, only 237 newborns were examined for transgenesis.

Colour, intensity and extent of coat pigmentation of the positive F_0 individuals varied over a wide range, ranging from a small, grey splotch on the back (about 5% of the coat without any visible eye melanization and any detectible transgenesis in skin probes outside of this region) to complete and homogenous, dark coat melanization. It was not possible to determine the exact number of animals showing phenotypic coat mosaicism owing to the appearence of natural, transgene-independent, rabbit coat patterns containing unpigmented coat regions (e.g., 'Dutch Belted'), previously masked by the albino phenotype of the rabbits. These rabbit coat patterns might be caused by modifications of loci controlling the pigmentation in the presence of an active tyrosinase. In pigmented mice, white coat regions were found to harbour no melanocytes or undifferentiated melanocytes in the skin (Takeuchi *et al.*, 1988; Bradl *et al.*, 1991; Mintz and Bradl, 1991).

To establish transgenic lines, four founders were chosen showing complete coat pigmentation (No. 2, 3 and 4) or the 'Dutch Belted' pattern (No. 5) with homogenous melanization of the pigmented coat regions (Fig. 1) and melanization of the eyes. This was based on our observations made in mice with the same gene construct indicating that the extent of coat pigmentation of the founders reflects the germline participation in the integration of the transgene (Aigner and Brem, 1994). A phenotypic coat mosaic (No. 1: map-like extent of the pigmentation over the whole body without sharp borders to the unpigmented regions) was bred as a control, expecting a lower percentage of transmission of the transgene to the offspring. The results agreed with our data derived from the ptrTyr4-transgenic founder mice. The classification according to the Sign test after Dixon and Mood (5%-level) (Sachs, 1992) revealed three of the four completely melanized animals as hemizygous (No. 2, 3 and 5) and the other full pigmented founder (No. 4) harbouring more than one integration site of the transgene, whereas the transgenic founder with the phenotypic coat mosaic (No. 1) showed germline mosaicism (Table 2).

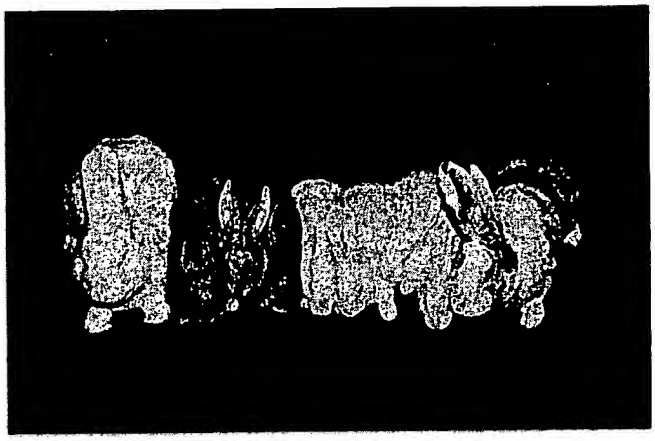


Fig. 1. Founder No. 5 (right, 'Dutch Belted') mated to the non-transgenic $ZIKA^R$ albino rabbit (left) with the F_1 offspring; the three hemizygous F_1 littermates show the same intensity of coat pigmentation.

Expression of the gene construct ptrTyr4

The examination of various tissue probes of an adult hemizygous individual (line 5) showing phenotypically complete and homogenous melanization of coat and pigment cell layers of the eyes by RT-PCR detected the expected expression of the murine gene construct in the pigmented tissues. Transgene expression was also observed in pancreas and testis (Fig. 2). In mice as well, expression of the same gene construct was found in tissues other than those expected (Beermann *et al.*, 1990). This might be caused by the presence of misdirected melanocytes in these tissues, which is found in various species. In addition to these tissues, in other tissue probes (liver and spleen) amplified cDNA derived from mRNA molecules with the small T intron of the SV40 cassette not spliced out was detected. Northern analysis of the same tissue probes showed no detectable bands indicating a low number of transcripts in the positive probes examined. Furthermore, the low number of melanocytes in the skin and different points of examination were suggested to be

Table 2. Production of transgenic lines by breeding transgenic founders to non-transgenic albino rabbits

Founder No.	*phenotype*	*F_1 generation*			*Classification of the founders according to the Sign test*
		total	*transgenic*	%	
1	coat mosaic	25	7	28	germline mosaic
2	completely pigm.	21	10	48	hemizygous
3	completely pigm.	15	7	47	hemizygous
4	completely pigm.	35	26	74	2 independent integration sites
5	'Dutch Belted'	7	3	43	hemizygous
Total: $n = 5$		103	53	–	–

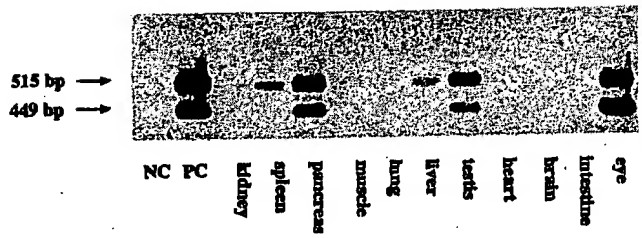


Fig. 2. RT-PCR analysis of various tissue probes of a hemizygous rabbit (line 5). The 1 kb SV40 fragment (*Bam*HI/*Sal*I) was used as a specific probe (Gorman *et al.*, 1982). In addition to the eye, the specific 449 bp signal also appeared in gonads and pancreas. Skin of a non-transgenic mouse was used as negative control (lane NC), pigmented skin of a ptrTyr4-transgenic mouse as positive control (lane PC). Besides these tissues, in liver and spleen amplified cDNA (515 bp) was also detected deriving from mRNA molecules with the 66 bp small T intron of the SV40 cassette not spliced out.

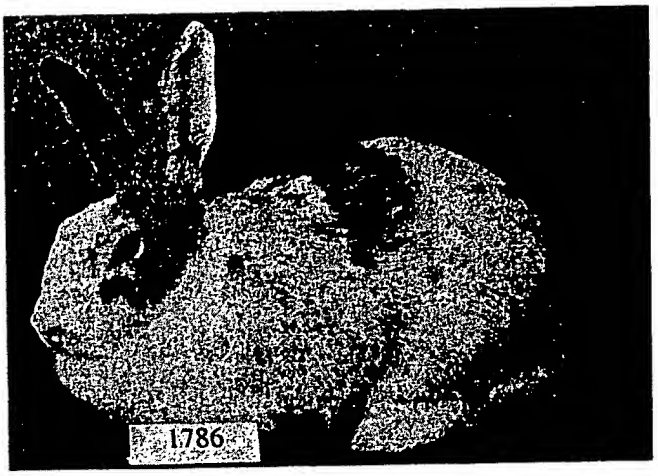


Fig. 3. Hemizygous F_1 rabbit with phenotypic coat mosaicism (line 4A). The phenotype was caused by overlapping of various rabbit coat patterns (such as 'Dutch Belted' and 'Japanese Pattern') indicating a high diversity of loci controlling pigmentation in ZIKA^R hybrid rabbits. In the transgenic rabbits the extent of macroscopically detectible melanization of the pigment cell layers of the eyes varied in a wide range.

responsible for varying results by different groups which have tried to detect the tyrosinase gene expression by northern analysis in the coat of pigmented mice (Müller *et al.*, 1988; Takeuchi *et al.*, 1988; Terao *et al.*, 1989).

Phenotypic examination of different foetal stages showed that tyrosinase-transgenic foetuses can be differentiated from non-transgenic ones already in the mid of gestation (17 days post conceptionem) by the melanization of the pigment cell layers of the eyes (data not shown). Thus, the temporal regulation of tyrosinase during embryonic development, as observed in mice (Beermann *et al.*, 1992), was also rescued in transgenic rabbits.

Correlation between genotype and phenotype in generations F_1 and F_2

A prerequisite for the use of tyrosinase as marker gene is the stable germline transmission and expression of the integrated transgene copies in the following generations. In the F_1 and F_2 generation, a strict correlation between coat pigmentation and detection of the transgene by means of molecular genetic methods was found, i.e. only the pigmented offspring ($n = 95$) inherited the transgene, whereas none of the albino littermates examined ($n = 62$) was detected as transgenic. The transgenic rabbit lines were designated according to the number of the founder animal listed in Table 2. In the F_1 and F_2 generation, transgenic individuals with one integration site of the gene construct appeared to harbour phenotypic coat mosaicism (non-homogeneous melanization with coat regions coloured in different tones and shades). This made it difficult to analyse the correlation between phenotype and genotype of the transgenics (Fig. 3). As observed in mice, the intensity and heterogeneity of coat pigmentation increased with the age of the individuals.

In general ptrTyr4-transgenic F_1 offspring showed a greater variety in hue and extent within the lines compared to the mice transgenic for the same gene construct (Aigner and Brem, 1994). However, the intensity of coat melanization differed only within proportionally confined limits in the lines 1, 2, 3 and 5 harbouring one integration site (phenotypically examined rabbits: line 1: $n = 5$; line 2: $n = 7$; line 3: $n = 4$; line 5: $n = 3$ (Fig. 1)). Line 4 derived from the transgenic founder with two independent integration sites of the gene construct. The transgenic offspring divided by phenotype into four different groups of coat colour intensity (group a–d: $n = 20$). By Southern analysis, the two less intense coloured groups a and b ($n = 11$) were shown to be linked with a reduced signal pattern compared to their darker pigmented siblings of groups c and d ($n = 9$). The reduced copy number of group a and b was confirmed by slot blot analysis and semiquantitative PCR (data not shown). Following the hybridization signal pattern, line 4 split in two lines 4A (with reduced copy number) and 4B. The strict correlation between copy number and expression of the gene construct in the F_1 generation of the transgenic lines was also observed in mice transgenic for the tyrosinase (Aigner and Brem, 1994). Comparison of the transcription levels of the endogenous gene and the transgene was not performed.

The estimation of the copy number by slot blot hybridization revealed about 10 copies of the transgene per cell in hemizygous animals of lines 1, 2 and 3, whereas line 4A harboured approximately 30 copies, line

5 about 60 copies and line 4B 100 copies (data not shown). Different coat colours occured in the differing lines making it difficult to classify the lines by the intensity of coat melanization.

For the production of homozygous animals, two phenotypically similar, hemizygous siblings of the F_1 generation were bred. Fifty-four offspring were born, of which 42 (78%) were pigmented. A strict copy number dependent expression of the tyrosinase gene construct, as seen in mice, was expected to result in two different levels of coat colour intensity: darker pigmented homozygous and lighter melanized hemizygous siblings (Aigner and Brem, 1994). Fig. 4 shows a litter of the F_2 generation, where the phenotypic classification and the differentiation of the transgenic F_2 rabbits in homozygous and hemizygous individuals by semiquantitative PCR matched completely. For other F_2 siblings however, the intensity of coat pigmentation did not always reflect the transgenic status of the animals (data not shown). Moreover, different coat colours and more than two levels of coat colour intensities in the same litter appeared. These deviations of the intensity of coat colour in relation to the genotype within the transgenic lines might not be caused by a copy-number-independent expression of the transgene, but by the ZIKA^R hybrid strain used. After rescuing the pigmentation with the murine wild-type tyrosinase gene construct, the high variety in the phenotype of individuals, as seen in mice, indicated a large diversity in the loci controlling the melanization, which was previously masked in the albino phenotype. The use of a heterologous gene construct seems to play a minor role for the deviations described. Thus, tyrosinase gene transfer in inbred rabbit strains might increase the correspondence between phenotype and genotype of the transgenics by making it easier to relate the intensity of coat pigmentation to the copy number of the animals.

Stability of transmission

In total, 109 transgenic rabbits were examined by Southern analysis for the stability of transmission of the gene construct, including hemizygous rabbits derived from backcrosses of hemizygous F_1 animals to non-transgenic animals ($n = 14$) and the transgenic animals of the F_1 and F_2 generation ($n = 95$). One case of deletion of transgene copies without detectible rearrangements was observed in the F_2 generation of line 5 (data not shown). Compared to the percentage of irregularities of transmission in ptrTyr4-transgenic mice (F_1 and F_2: 0.6%, $n = 509$), the result in rabbits (0.9%) was similar.

Conclusions

In this study we examined tyrosinase as a marker for transgenesis in rabbits by analysing the correlation between phenotype and genotype of tyrosinase-transgenic rabbits. The value of tyrosinase as a marker gene in rabbits was shown by the strict correlation between integration and expression of the transgene and the stable germline transmission of the integrated gene construct according to the Mendelian pattern of inheritance. In addition, a correlation between the intensity of coat pigmentation and the number of integrated transgene copies was shown in the F_1 generation. Although irregularities in coat pigmentation limited the use of tyrosinase for the identification of homozygous individuals in the

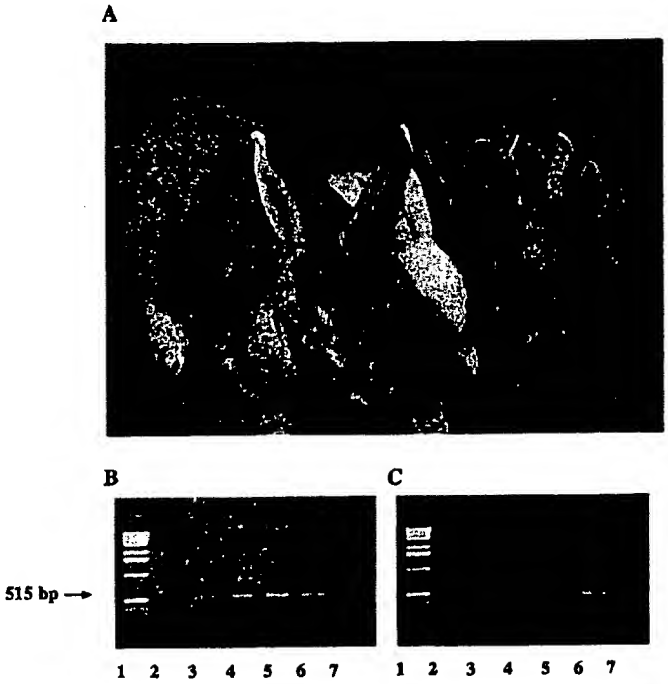


Fig. 4. (A) Generation of homozygous rabbits: F_2 offspring of two hemizygous littermates (line 5) with strict correlation of phenotype and genotype. One hemizygous parent is shown with one non-transgenic albino offspring, two hemizygous (same intensity of coat colour) and one homozygous (darker coloured) offspring. (B) Analysis by semiquantitative PCR revealed a double copy number of the transgene per cell in the darker coloured offspring (lane 5) compared to its hemizygous siblings (lanes 3, 4) and parents (lane 7). As a positive control, double the amount of DNA of the hemizygous parent was used (lane 6), giving rise to a signal intensity comparable to that of the homozygous offspring. DNA of an albino sibling was used as a negative control (lane 2); the molecular weight marker is the 1 kb ladder (Life Technologies, Eggenstein, Germany) (lane 1). (C) The control of the amount of DNA used in the semiquantitative PCR was carried out by semiquantitative PCR for the endogenous rabbit transferrin gene. The application of the probes corresponds to Fig. 4B. As positive control the double amount of DNA of the hemizygous parent was used (lane 6); the negative control was mouse genomic DNA (lane 2). The same signal intensities of the individuals examined show that the same amount of DNA was used in the semiquantitative PCR.

ZIKA^R hybrid strain used, breeding control of transgenic rabbit lines was facilitated by co-transfer and co-integration of the murine wild-type tyrosinase gene construct in accord with previous observations made in transgenic mice (Beermann *et al.*, 1991; Overbeek *et al.*, 1991). Thus, in line 5, co-inheritance of a co-integrated gene construct causing tissue specific production of IGF-1 (insulin-like growth factor 1) in the mammary gland (Zinovieva *et al.*, manuscript in preparation) was observed in 7 generations. The use of rabbits in gene transfer experiments may be required as alternative species to mice for certain applications (Weidle *et al.*, 1991). Furthermore, rabbits combine a short generation interval and low costs for animal husbandry with use as farm animals which prepares them for experiments in genetic engineering, e.g. production of foreign proteins in the mammary gland (Brem *et al.*, 1994).

Acknowledgements

We thank Prof. G. Schütz, Abteilung Molekularbiologie der Zelle I des DKFZ Heidelberg, for kindly providing the gene construct ptrTyr4, U. Huffstadt for the microinjection of the transgene and Dr M. Müller for critically reading the manuscript.

References

Aigner, B. and Brem, G. (1993) Tyrosinase as marker gene and model for screening transgenes in mice and rabbits. *Theriogenology* **39**, 177.

Aigner, B. and Brem, G. (1994) Tyrosinase gene as a marker gene for studying transmission and expression of transgenes in mice. *Transgenics* **1**, 417–29.

Aigner, B. and Brem, G. (1995) Detection of homozygous individuals in gene transfer experiments by semiquantitative PCR. *BioTechniques* **18**, 574–8.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons.

Banfield, D.K., Chow, B.K.C., Funk, W.D., Robertson, K.A., Umelas, T.M., Woodworth, R.C. and MacGillivray, R.T.A. (1991) The sequence of rabbit liver transferrin cDNA. *Biochim. Biophys. Acta* **1089**, 262–5.

Beermann, F., Ruppert, S., Hummler, E., Bosch, F.X., Müller, G., Rüther, U. and Schütz, G. (1990) Rescue of the albino phenotype by introduction of a functional tyrosinase gene into mice. *EMBO J.* **9**, 2819–26.

Beermann, F., Ruppert, S., Hummler, E. and Schütz, G. (1991) Tyrosinase as a marker for transgenic mice. *Nucl. Acids Res.* **19**, 958.

Beermann, F., Schmid, E. and Schütz, G. (1992) Expression of the mouse tyrosinase gene during embryonic development: recapitulation of the temporal regulation in transgenic mice. *Proc. Natl Acad. Sci. USA* **89**, 2809–13.

Besenfelder, U. and Brem, G. (1993) Laparoscopic embryo transfer in rabbits. *J. Repr. Fertil.* **99**, 53–6.

Bradl, M., Larue, L. and Mintz, B. (1991) Clonal coat color variation due to a transforming gene expressed in melanocytes of transgenic mice. *Proc. Natl Acad. Sci. USA* **88**, 6447–51.

Brem, G. (1993) Transgenic animals. In Rehm, H.J. and Reed, G., eds., *Biotechnology – Volume 2: Genetic Fundamentals and Genetic Engineering*, pp. 745–832. Weinheim: VCH.

Brem, G., Hartl, P., Besenfelder, U., Wolf, E., Zinovieva, N. and Pfaller, R. (1994) Expression of synthetic cDNA sequences encoding human insulin-like growth factor-1 (IGF-1) in the mammary gland of transgenic rabbits. *Gene* **149**, 351–5.

Chomcynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–9.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**, 1044–51.

Halaban, R., Moellmann, G., Tamura, A., Kwon, B.S., Kuklinska, B., Pomerantz, H.P. and Lerner, A.B. (1988) Tyrosinase of murine melanocytes with mutations at the albino locus. *Proc. Natl Acad. Sci. USA* **85**, 7241–5.

Hearing, V.J. (1987) Mammalian monophenol monooxygenase (tyrosinase): purification, properties and reactions catalyzed. In Kaufman, S., ed., *Methods in Enzymology – Metabolism of Aromatic Amino Acids and Amines*, pp. 154–65. New York: Academic Press.

Hearing, V.J. and Tsukamoto, K. (1991) Enzymatic control of pigmentation in mammals. *FASEB J.* **5**, 2902–9.

Hogan, B., Costantini, F. and Lacy, E. (1986) *Manipulating the Mouse Embryo*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (1990) *PCR protocols: a Guide to Methods and Applications*. San Diego, CA: Academic Press Inc.

Jackson, I.J. and Bennett, D.C. (1990) Identification of the albino mutation of mouse tyrosinase by analysis of an in vitro revertant. *Proc. Natl Acad. Sci. USA* **87**, 7010–4.

Kwon, B.S., Halaban, R. and Chintamaneni, C. (1989) Molecular basis of mouse himalayan mutation. *Biochem. Biophys. Res. Commun.* **161**, 252–60.

Mintz, B. and Bradl, M. (1991) Mosaic expression of a tyrosinase fusion gene in albino mice yields a heritable stripped coat color pattern in transgenic homozygotes. *Proc. Natl Acad. Sci. USA* **88**, 9643–7.

Müller, G., Ruppert, S., Schmid, E. and Schütz, G. (1988) Functional analysis of alternatively spliced tyrosinase gene transcripts. *EMBO J.* **7**, 2723–30.

Overbeek, P.A., Aguilar-Cordova, E., Hanten, G., Schaffner, D.L., Patel, P., Lebovitz, R.M. and Lieberman, M.W. (1991) Coinjection strategy for visual identification of transgenic mice. *Transgenic Res.* **1**, 31–7.

Ruppert, S., Müller, G., Kwon, B. and Schütz, G. (1988) Multiple transcripts of the mouse tyrosinase gene are generated by alternative splicing. *EMBO J.* **7**, 2715–22.

Sachs, L. (1992) *Angewandte Statistik*. New York: Springer Verlag.

Shibahara, S., Okinaga, S., Tomita, Y., Takeda, A., Yamamoto, H., Sato, M. and Takeuchi, T. (1990) A point mutation in the tyrosinase gene of BALB/c albino mouse causing the cysteine → serine substitution at position 85. *Eur. J. Biochem.* **189**, 455–61.

Silvers, W.K. (1979) *The Coat Color of Mice – A Model for*

Mammalian Gene Action and Interaction. New York: Springer Verlag.

Takeda, A., Matsunaga, J., Tomita, Y., Tagami, H. and Shibahara, S. (1989) Molecular analysis of the DNA segments cross-hybridizable to the tyrosinase gene in patients afflicted with oculocutaneous albinism. *Tohoku J. Exp. Med.* **159**, 333–40.

Takeuchi, S., Yamamoto, H. and Takeuchi, T. (1988) Expression of tyrosinase gene in amelanotic mutant mice. *Biochem. Biophys. Res. Commun.* **155**, 470–5.

Tanaka, S., Yamamoto, H., Takeuchi, S. and Takeuchi, T. (1990) Melanization in albino mice transformed by introducing cloned mouse tyrosinase gene. *Development* **108**, 223–7.

Terao, M., Tabe, L., Garattini, E., Sartori, D., Studer, M. and Mintz, B. (1989) Isolation and characterization of variant cDNAs encoding mouse tyrosinase. *Biochem. Biophys. Res. Commun.* **159**, 848–53.

Weidle, U.H., Lenz, H. and Brem, G. (1991) Genes encoding a mouse monoclonal antibody are expressed in transgenic mice, rabbits and pigs. *Gene* **98**, 185–91.

Yamamoto, H., Takeuchi, S., Kudo, T., Sato, C. and Takeuchi, T. (1989) Melanin production in cultured albino melanocytes transfected with mouse tyrosinase cDNA. *Jpn J. Gen.* **64**, 121–35.

Yokoyama, T., Silversides, D.W., Waymire, K.G., Kwon, B.S., Takeuchi, T. and Overbeek, P.A. (1990) Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice. *Nucl. Acids Res.* **18**, 7293–8.

Thrombosis and Haemostasis © F. K. Schattauer Verlagsgesellschaft mbH (Stuttgart) 78(1) 537–542 (1997)

Current Progress in the Production of Recombinant Human Fibrinogen in the Milk of Transgenic Animals

Stephen P. Butler[1], Kevin van Cott[2], Anuradha Subrumanian, Francis C. Gwazduaskas[1] and William H. Velander[2]

[1]Department of Dairy Science, Virginia Polytechnic Institute and State University, Blacksburg, VA; [2]Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Introduction

Transgenic livestock as an alternative source of recombinant fibrinogen

Many applications for fibrin sealant (FS) have been developed over past two decades due to the increased availability of human fibrinogen (hfib) available from cryoprecipitation processing of human plasma (1). In addition to use as an hemostatic agent, new FS-applications include carrier matrix applications for the delivery of drugs and biologics (1–5). Estimates of annual US clinical need are now greatly in excess of the approximately 300 kg per year which can be harvested using the currently low yield, cryoprecipitation methods on the total of 7 million liters of human plasma available for fractionation in the US (6). Thus, the recombinant source must be capable of economically supplying about 1000 kg/yr in order to replace the maximum possible hfib obtainable from current plasma fractionation by cryoprecipitation. The use of the mammary gland of transgenic livestock as a bioreactor for producing rhfib is currently being evaluated in dairy livestock and pigs. While dairy cattle are the most prodigious milk producers, the goat, sheep, and pig have all demonstrated the ability to produce g/l levels of recombinant proteins. The annual milk yields of these livestock are about 1000 liters per year for the goat (7), 500 l/yr for the sheep, and 100 to 300 l/yr for the pig. Thus, only 2000 sows would be needed to produce 1000 kg/yr or more rhfib at a concentration of 5 g/l in milk. Using the transgenic mammary gland for production of rhfib has several necessary requirements. First, DNA constructs which use mammary specific promoters effective for expressing cDNA, minigene, or genomic fibrinogen coding sequences must be formulated. These constructs are then introduced to embryonic cells, typically using microinjection into the pronucleus of zygotes. Second, the cointegration of constructs containing each of the α, β, and γ encoding sequences for hfib into a transcriptionally responsive chromosomal domain is needed. For the purposes of pharmaceutical production at large-scale, a single cointegration site and associated transgene copy number in that loci is desirable for the facile establishment of a phenotypical and genotypical stable lineage (8). Mosiacism and multiple integration sites frequently occurs in founder animals and this complicates analysis of founder animals (9). Thus, phenotype and genotype can not be reliably defined in transgenic animals until successive generations of offspring obtained from outbreeding with nontransgenic animals are analyzed (8–9).

Regulatory motifs using murine whey acidic protein and ovine β-lactoglobulin promoters to express rhfib

Three different promoters, ovine β-lactoglobulin (oBLG; 10,11) and two versions of the murine whey acidic protein (mWAP; reviewed in 12, W. Velander, unpublished data) have been effective in achieving the co- expression of each rhfib chain at high levels relative to that obtained in cell culture. Both oBLG and mWAP are whey proteins which naturally occur in milk at about 1-2 g/l in the milk of sheep and mice, respectively. However, the levels of recombinant proteins obtained using regulatory elements from the mWAP and oBLG genes have widely ranged from 0.001 to 30 g/l in the milk of different transgenic animal species (reviewed in 12). The structure of the transgene can be used to achieve different levels of expression in the milk of transgenic animals. Hence, the ability to target different levels of expression can be used to optimize the yields of fully functional product relative to limitations encountered in post translational modifications and the secretory pathway (13). To that end, the mWAP promoter has been one of the most effective regulatory elements across species for targeting a relatively narrow range of mammary specific expression range by the use of a specific mWAP-transgene motif (8,13).

Fig.1 shows the structure of oBLG and mWAP hfib-DNA constructs which have been used in efforts to express rhfib in

Correspondence to: W.H. Velander, 142 Randolph Hall, Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0211, USA, Tel: +1-540-2317869; Fax: +1-540-2315022; E-mail: wvelande@vt.edu

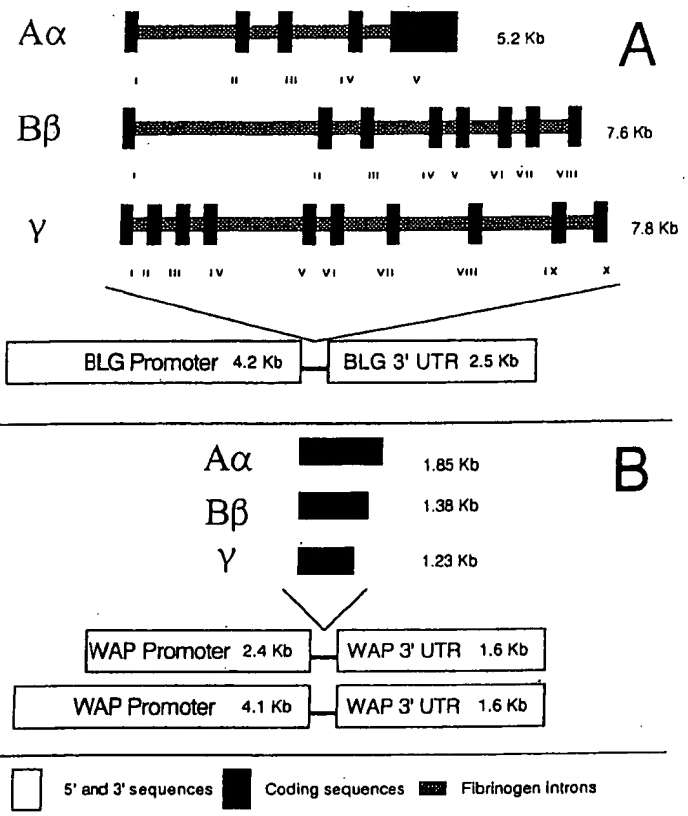


Fig. 1. Examples of human fibrinogen transgene design. A. The genomic regions for each fibrinogen chain (Aα,Bβ,γ) were inserted between 4.2 Kb of the ovine Beta lactoglobulin (BLG) promoter and 2.5 Kb of the BLG associated 3' untranslated region (3'UTR) to produced three individual constructs (adapted from 10). B. The corresponding cDNAs for each fibrinogen chain were inserted between either 2.4 or 4.1 Kb of the murine whey acid protein (WAP) promoter and 1.6 Kb of the WAP associated 3'UTR (Velander , unpublished data).

the milk of transgenic mice, pigs and sheep. A general cloning strategy used in fibrinogen transgene design incorporates inserting a coding sequence of DNA for each individual chain between the milk gene promoter and its associated 3' untransl-ated region (3'UTR). In the examples reviewed here, a total of three separate constructs have been used to express the mature rhfib in the milk. Fig. 1A shows the transgene construct layout for the oBLG driven genomic fibrinogen chains (adapted from 10). The genomic sequences coding for each chain were trimmed of noncoding 5' and 3' flanking sequences and then inserted between the oBLG promoter and 3'UTR DNA sequences. Thus, transcription initiation is regulated by the oBLG promoter with transcription termination and polyad-enylation being controlled by the oBLG-3'UTR. The mWAP transgenes illustrated in fig.1B follow a similar format and have been successful at expressing cDNAs -reviewed in (12).

Expression of rhfib in milk of mouse and pig using different mWAP promoters

The contrast in expression levels obtained in different species has been previously established by results obtained using the 2.4 kb mWAP promoter linked to the cDNA sequence of human protein C (hPC). Transgenic mice having this construct that secreted rhPC in milk at only 0.001 to 0.010 g/l -reviewed in (12). The expression level of rhfib in mouse milk using the same 2.4 kb mWAP format ranged from 0.01 to 0.05 g/l for completely assembled fibrinogen.

One monogenic transgenic mouse line containing only the 2.4 kb mWAP-β-chain-hfib-cDNA construct expressed β- chain rhfib alone at about 0.03 g/l. In contrast to the 2.4 kb mWAP promoter, a longer promoter using 4.1 kb of mWAP and the cDNA of hPC produced recombinant hPC (rhPC) at 0.5 to 1 g/l in the milk of transgenic mice (W. Velander, unpublished data). An expression level of 0.1 to 0.6 g/l rhfib has been obtained in three founder animals having the 4.1 kb mWAP-hfib-cDNA. These observations suggest that the longer mWAP promoter may be able to express cDNAs to relatively high levels in milk of livestock. The two promoters were derived from different genomic mouse libraries and upon sequence analysis, it was determined that a high degree of homology is present in the proximal 1.5 kb with sequence divergence in their distal regions (14). Since both the 2.4 kb and 4.1 kb promoters share a high degree of homology in the proximal region from the transcriptional start, the ability of the longer promoter to express cDNAs at high levels may be due the presence of enhancer elements and/or the lack of repressor elements found in distal region as suggested in (15).

Fig 2A. Shows a western blot analysis of nonreduced SDS PAGE (4 to 12% gradient) analysis of milk from 2.4 kb mWAP-hfib-cDNA transgenic mice. These milks showed a rhfib species of about 340 kDa which was similar to that

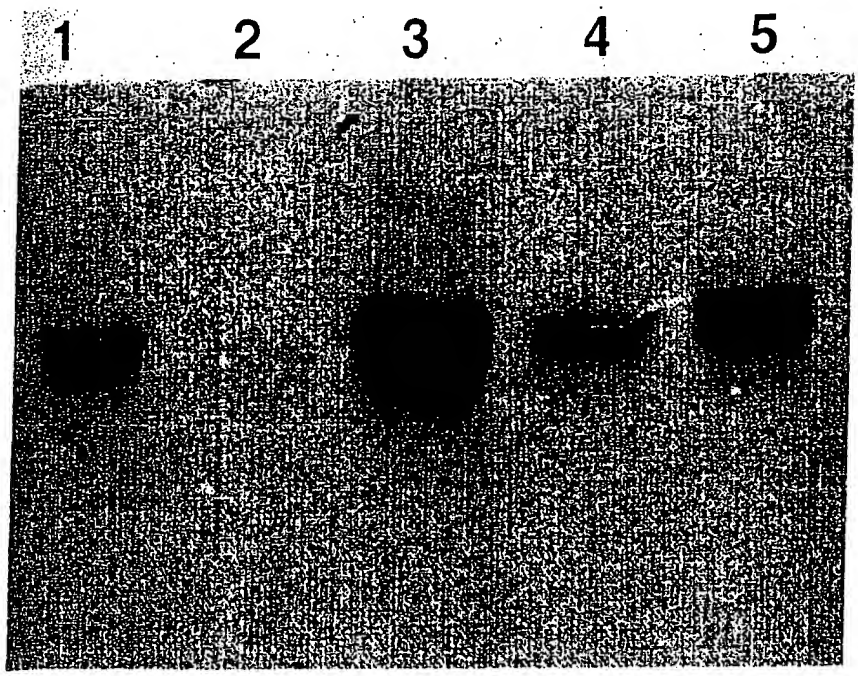


Fig. 2A. Western analysis of partially purified pooled milk samples from 2.4 mWAP-hfib-cDNA, αβγF1 transgenic mice. Transgenic, nontransgenic, and hfib-spiked nontransgenic milk samples were pre-treated by DEAE ion exchange chromatography, followed by $Zn2^+$ precipitation at 8 mM $ZnCl_2$ to selectively remove mouse fibrinogen from the milk (Velander, unpublished data). Mouse fibrinogen partitions into the 8 mM ZnCl2 precipitate, while rhfib remains in the supernatant. Sa-mples were separated on a 4-12% SDS PAGE under non-reducing conditions, transferred to a nitrocellulose membrane, probed with a polyclonal anti-hfib (Celsus, Cincinnati, OH), and visualized by metal enhanced DAB staining (Pierce, Rockford, IL). Lane 1: hfib standard, purified from human plasma, 50 ng. Lane 2: 8mM $ZnCl_2$ supernant from nontransgenic mouse milk, 10 μg total protein loaded. Lane 3: 8mM Zn Cl_2 supernant of nontransgenic mouse milk spiked with 100 ng of hfib. Lane 4: 8mM $ZnCl_2$ supernant of transgenic mouse milk (50 ng rhfib loaded). Lane 5: hfib standard, 100 ng.

obtained for hfib. Fig 2B shows western analyses using a polyclonal antibody detection of SDS-PAGE under reducing conditions for several different F1 offspring from outbreeding of a 4.1 kb-mWAP-hfib-cDNA transgenic founder. These transgenic mice produced about 0.1 to 0.3 g/l rhfib in their milk as detected by polyclonal ELISA. As can be seen in Fig 2B, the expression levels of rhfib are relatively stable throughout lactation. Purified rhfib from the milk of mice containing mWAP constructs was shown to be a functional substrate for dual treatment by both human thrombin and FXIII as clots were obtained and cross-linked fragments appeared in western analysis of the reduced clots (data not shown). Milk from rhfib produced by transgenic mice containing the oBLG-hfib- gene also developed cross-linked fibrin clots after treatment of the milk by human thrombin in the presence of FXIII (10).

Our initial studies in expressing rhfib in the porcine mammary gland have been with cDNA constructs containing the 2.4 kb mWAP promoter. Pronuclear microinjection of equilmolar concentrations of α, β, and γ hfib constructs into one-celled embryos has resulted in 3 litters containing 10 piglets. About 320 microinjected embryos were transfered into a total of 8 synchronized surrogate mothers (40 embryos per recipient). Five of the 10 total piglets born contained at least two of the mWAP-hfib transgenes. Only one of these piglets contained all three transgenes in DNA isolated from tail tissue samples, and it was the only founder animal that contained hfib- transgenes in the germline. Multiple loci were detected in offspring obtained by the outbreeding of this gilt with a nontransgenic boar. Nine of 11 of these offspring contained at least two or three of the α-,β-, and γ-hfib transgenes. Only alpha chain was found in the milk of the founder αβγ-hfib transgenic pig where somatic tissue mosaicism or disfunctional transgene insertion sites may have occurred for some of these transgene insertion sites. Milk from offspring having all three hfib-transgenes will be analyzed to determine if somatic tissue mosiacism or dysfunctional integration sites is occurring in the mammary gland of this single founder animal. Thus, co-integration of separate, multiple transgenes can introduce an increased level of complexity to making and time frame needed for evaluating transgenic livestock.

Expression of rhfib in milk of mouse and sheep using the oBLG promoter

The differences of expression obtained for the 4.2 oBLG promoter using cDNA or gene sequences in mice and sheep are considerable and different than those obtained using mWAP driven constructs in mice and pigs. In contrast to mWAP promoter in mice and pigs, the use of the oBLG promoter to give high levels of expression in the milk of mice and sheep apparently requires genomic coding sequences. Thus, oBLG-hfib-gene constructs were selected in experiments to express rhfib in the milk of transgenic sheep. As also occurred in transgenic mice containing the mWAP-hfib- cDNA constructs, about 85% of oBLG-hfib-gene transgenic mice contained all three hfib-transgenes (10). Transgene copy numbers varied among the different lines but the ratio of the three constructs within a given locus was usually equal. The milk secretion levels observed varied widely from 0.03 to 2 g/l with the average across founding lines being 0.6 g/l. The

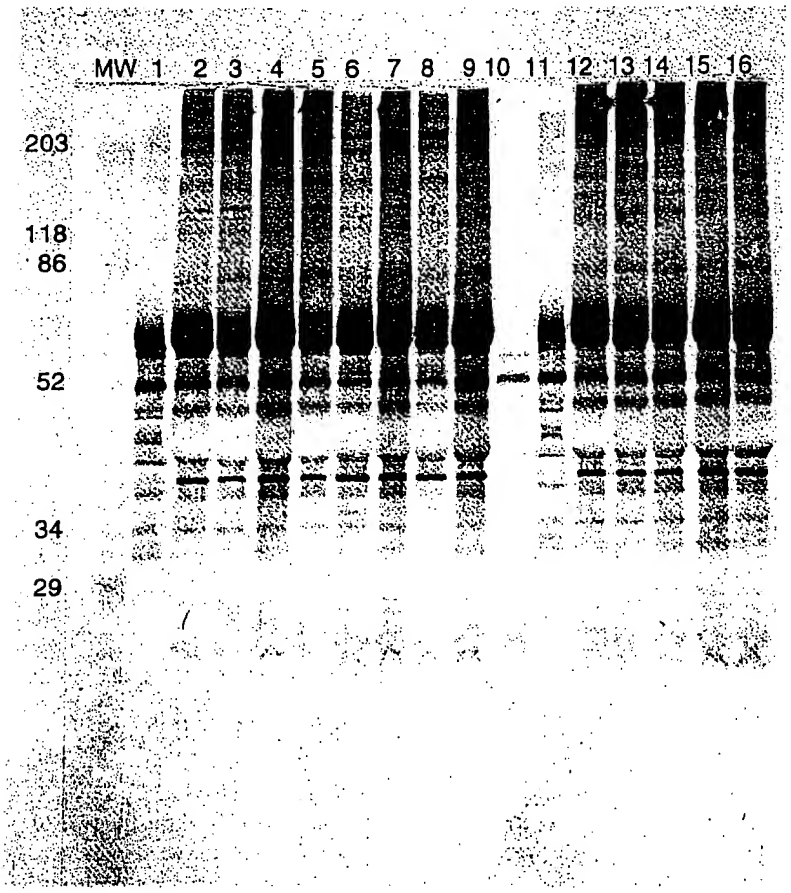


Fig. 2B. Western analysis of daily milk samples from the first (Lanes 2-9) and second (Lanes 12-16) lactations of 4.1 mWAP-hfib-cDNA, αβγ F1 transgenic mice. Equal volumes of defatted milk from transgenic and nontransgenic were applied directly to the gel. Milk samples were electrophoresed under reducing conditions, transferred to a PVDP membrane, probed with a polyclonal antibody to hfib (Celsus, Cincinnati, OH) and visualized with metal enhanced DAB substrate (Pierce, Rockford, IL). Lane 1: hfib standard purified from plasma (100 ng). Lane 2: milk from mouse 10-23, Day 6. Lane 3: milk from mouse 10-23, Day 12. Lane 4: milk from mouse 10-28, Day 7. Lane 5: milk from mouse 10-28, Day 11. Lane 6: milk from mouse 10-13, Day 7. Lane 7: milk from mouse 10-13, Day 11. Lane 8: milk from mouse 10-27, Day 6. Lane 9: milk from mouse 10-27, Day 10. Lane 10: Nontransgenic mouse milk. Lane 11: hfib (100 ng). Lane 12: milk from mouse 10-23, Day 5. Lane 13: milk from mouse 10-28, Day 6. Lane 14: milk from mouse 10-13, Day 5. Lane 15: 10-13, Day 5. Lane 16: 10-27, Day 5.

average expression level for outbred F1 transgenic mice containing the mWAP- hfib-cDNA mice was about 0.3 g/l for 3 lines expressing rhfib at detectable levels. The total number of integrated copies did not correlate with the amount of rhfib secreted, indicating that the constructs are more likely influenced by chromosomal positioning which is commonly observed with transgenes. Pronuclear microinjection of sheep zygotes using the oBLG-hfib genomic constructs resulted in the generation of 9 founding animals (11).

Expression data from four ewes has been reported where 3 animals contained the α-,β-, and γ-transgenes in their genome. A fourth animal contained only genes for the beta and gamma chains. Upon hormone induced lactation at 4 months of age, milk from two of the trigenic animals contained 0.5 g/l and a third ewe contained about 5 g/l. Purified material from the highest expressing ewe was subjected to amino terminal analysis before and after thrombin cleavage resulting in the expected sequences for all three chains. SDS-PAGE analysis under reducing conditions demonstrated that the recombinant fibrinogen can undergo cross-linking to form γ-γ dimers after dual treatment by human thrombin and factor XIII .

Comparisons to rhfib production in cell culture

The production of biologically active hfib requires coordinated expression of three separate fibrinogen genes so that co-translation and assembly of a complex hexameric structure occurs. Transgenic mice containing all three transgenes have secreted a wide range of partially assembled rhfib. For both mWAP-hfib-cDNA and sBLG-hfib-gene transgenes, an estimated range of 10 to 100% of rhfib chains were fully assembled in milks of different transgenic mouse lineages (10, W. Velander.unpublished data). Thus far, no correlation has been established with gene copy number or αβγ transgene ratio and rhfib assembly. Assembly phenomena has been extensively studied in the production of rhfib by mammalian cells in culture (16-19). The mechanism for assembly apparently requires a pool of precursor fibrinogen polypeptides to be retained intracellularly by chaperone proteins which occur within the ER and Golgi secretory apparatus. The cDNAs of human fibrinogen have been expressed in transfected monkey kidney fibroblast (COS1; 17,20), baby hamster kidney (BHK; 19), chinese hamster ovary (CHO; 21) and Hep G2 cells (22). Reports of secretion of individual chains and chain complexes into the media have been limited to individual Aα chain and γ chains, an Aα-γ chain complex and half molecules (αβγ). Our work with monogenic mice containing only α or β cDNAs show that the mammary gland can secrete individual Aα chains and Bβ chains into milk, respectively. In one mosiac, bigenic founder pig containing the 2.4 mWAP α- and β-hFib-cDNA constructs, only the Aα chain polypeptide was secreted. In contrast to

the milk of trangenic mice, individual Bβ chains have not been found as secretion products from any single β-hfib-DNA transfected cell line studied (20). For example, biP chaperones precipitated by anti-biP antibodies from cellular lysates of COS-1 beta-chain-only transfectants contained Bβ chain. This Bβ chain contained a mannose rich, endo-H sensitive carbohydrate suggesting this chain is not transported to the Golgi. Thus, the ability of the mammary gland to secrete separate Bβ chain indicates that its chaperone system operates differently than cell lines which do not secrete Bβ chain alone. The secretory behavior of mammary epithelial cells with rhfib is consistent with the secretion of immature forms of other recombinant proteins into milk such as obtained for pro-rhPC (23). Synthesis rates of functionally active rhfib at 1 pg rhfib/cell/day or less were observed with cultured mammalian cells (18) and about 0.5 pg/cell/day at a cell density of 108 cells/ml using yeast (24). In contrast, density of mammary epithelia is about 10^9 cells/ml while secreting about 0.1 to 1 g/l/hr rhfib in mice and sheep. In summary, one of the advantages of producing rhfib in the milk of transgenic livestock is the 10 and 100 fold higher cell density of the mammary gland relative to yeast and mammalian cells in culture (25).

Summary

The mammary gland of transgenic animals has several advantages for production of heterologous proteins including a high cell density that results in high concentrations of secreted protein. While the mammary gland appears to be able to secrete fully assembled recombinant human fibrinogen (rhfib) at 0.1 to 5 g/l levels, some unassembled rhfib chains are also secreted. Presently, the relationship between unassembled rhfib and the coordinated translation of each nascent rhfib polypeptide in the mammary epithelia is unknown. The secretion of fully and partially assembled rhfib is widely variable among mammalian cell lines and where previously no cell line has been shown to secrete beta chain alone. We have observed that mammary epithelia can secrete Bβ chain into milk as well as immature forms of other recombinant proteins, suggesting it likely uses a different secretory pathway than does the liver. This difference in secretory behavior is possibly due to the natural design of milk, where the precise regulation of post translational modifications and intracellular pools of nascent polypeptides needed to achieve fibrinogen assembly may be less important to fulfill the nutritional function of most milk proteins.

References

1. Alving BM, Weinstein MJ, Finlayson JS, Menitove JE, Frantantoni JC. Fibrin sealant: summary of a conference on characteristics and clinical uses. Transfusion 1995; 35:783-790.
2. Greco F, de Palma L, Spagnolo N, Rossi A, Specchia N, Gigante A. Fibrin-antibiotic mixtures: an in vitro study assessing the possibility of using a biologic carrier for local drug delivery. J Biomed Mat Res 1991; 25:39-51.
3. MacPhee M, Campagna A, Kidd R, Best A, Drohan WN. Fibrin sealant as a delivery vehicle for sustained and controlled release of chemotherapy agents. In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
4. MacPhee M, Nunez H, Hennings R, Campagna A, Grummon G, Harding S, Drohan W. Sustained release in vitro and *in vivo* of antibiotics from fibrin sealant: In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
5. Lasa C, Hollinger J, Drohan W, MacPhee M. Bone induction by demineralized bone powder and partially purified osteogenin using a fibrin-sealant carrier. In: Surgical Adhesives and Sealants: Current Technology and Applications. Sierra D. and Saltz R.,eds. Technomic Publishing Co. Lancaster, PA 1996.
6. Paleyanda R, Young J, Velander, W *et al.* The expression of therapeutic proteins in transgenic animals. In Hoyer LW and Drohan WN, eds., *Recombinant Technology in Hemostasis and Thrombosis*.New York: Plenum Press. 1991:197-209.
7. Haenlein GFW. Dairy Goat industry of the United States. J dairy Sci 1981; 64: 1288-1304.
8. Van Cott KE, Lubon H, Russell CG, Butler SP, Gwazdauskas FC, Knight J, Drohan WN, Velander WH. Phenotypic and genotypic stability of multiple lines of transgenic pigs expressing recombinant human protein C. Transgenic Res 1997; 6:1-10.
9. Velander WH, Johnson JL, Page RL *et al.* High-level expression of a heterologous protein in the milk of transgenic swine using the cDNA encoding human protein C. Proc Natl Acad Sci USA 1992; 89:12003- 12007.
10. Punkard D, Cottingham I, Garner I, Bruce S, Dalrymple M, Lasser G, Bishop B, Foster D. High-level expression of recombinant human fibrinogen in the milk of transgenic mice. Bio/Technol 1996; 14:867-871.
11. Cottingham I, McKee C, Temperley S, Garner I, Prunkard D, Lasser G, Bishop P, Foster D. Human fibrinogen from the milk of transgenic sheep. In: Tissue Sealants: Current Practice, Future Uses. Cambridge Institute, Newton Upper Falls, MA. March 30 April 2 1996 (abstract).
12. Lubon H, Paleyanda RK, Velander WH *et al.* Blood proteins from tra-nsgenic animal bioreactors.Transfusion Med Rev 1996;10:131-143.
13. Van Cott KE, Williams BL, Velander WH *et al.* Affinity purification of biologically active and inactive forms of recombinant human protein C produced in porcine mammary gland. Journal of Molecular Recognition 1997; 9: 407-414.
14. Paleyanda RK, Zhang D, Hennighausen L, McKnight RA, Lubon H. Regulation of human protein C gene expression by the mouse WAP promoter. Transgenic Res.1994; 3: 335-343.
15. McKnight RA, Spencer M, Dittmer J, Brady JN, Wall RJ, Henninghausen L. An Ets site in the whey acidic protein gene promoter mediates transcriptional activation in the mammary gland of pregnant mice but is dispensable during lactation. Mol. Endocrinol. 1995; 9: 717-724.
16. Yu S, Sher B, Kudryk B, Redman CM. Fibrinogen precursors, order of assembly of fibrinogen chains, J Bio Chem 1984; 259:10574-10581.
17. Hartwig R and Danishefsky KJ. Studies on the assembly and secretion of fibrinogen, J Biol Chem 1991; 266:6578-6585.
18. Roy SN, Procyk R, Kudryk BJ, Redman CM. Assembly and secretion of recombinant human fibrinogen, J Biol Chem 1991; 266:4758-4763.

19. Huang S, Mulvihill ER, Farrell DH, Chung DW, Davie EW. Biosys-thesis of human fibrinogen. J Bio Chem 1993; 268: 8919-8926.
20. Danishefsky K, Hartwig R, Banerjee D, Redman C. Intracellular fate of fibrinogen Bb chain expressed in COS cells. Biochim Biophy Acta, 1990; 1048:201-208.
21. Lord ST, Strickland E, Jayjock E. Strategy for recombinant multichain protein synthesis: fibrinogen Bb-chain variants as thrombin substrates. Biochem 1996; 35:2342-2348.
22. Yu S, Sher B, Kudryk B, Redman CM. Intracellular assembly of human fibrinogen. J. Biol Chem 1983; 258:13407-13410.
23. Lee TK, Drohan WN, Lubon H. Proteolytic processing of human protein C in swine mammary gland. J Biochem 1995; 118:81-87.
24. Roy SN, Kudryk B, Redman CM. Secretion of biologically active recombinant fibrinogen by yeast. J Biol Chem 1995; 270: 23761-23767.
25. Morcol, T, Akers RM, Johnson J.L *et al*. The porcine mammary gland as a bioreactor for complex proteins. Annal NY Acad Sci 1994; 721:218-233.

The

Williar

Past

The use
product
as 1982
from a
physiol
into mo
the pr
integra
transmi
technol
possibl
over a
control
techniq
early e
to isola
of DN
develo
was sh
effects
large
1983,
had b
specifi
The
obviou
pathw
syster
genera
proble
injecti
viabil
transg
cows
Th
severa
prote
tissue
way
gland
milk
prote
(22,2

Corres
Way,
(301)

RESEARCH

GENETICALLY ENGINEERED FERTILE INDICA-RICE RECOVERED FROM PROTOPLASTS

Swapan K. Datta*, Alex Peterhans[1], Karabi Datta and Ingo Potrykus

Swiss Federal Institute for Technology, Plant Sciences, ETH-Zentrum, CH-8092, Zurich, Switzerland. [1]Present address: Rockefeller University, 1230 York Ave., NYC, NY 10021. *Corresponding author.

We have established an efficient protocol for plant regeneration from haploid Indica-type rice protoplasts. Incubation of these protoplasts with the selectable hygromycin phosphotransferase *(hph)* gene expressed under control of the 35S promoter of cauliflower mosaic virus (CaMV) and polyethyleneglycol (PEG), and subsequent culture in the presence of hygromycin B, led to the recovery of numerous resistant clones from which 77 plants were regenerated. Data from Southern analysis and enzyme assays proved that the transgene was stably integrated into the host genome and expressed, and that it was inherited in offspring.

Rice *(Oryza sativa)* is the world's most important crop plant[1]. Indica-type rice varieties feed more than 2 billion people, predominantly in developing countries[2]. Despite a variety of approaches, transgenic cereals can be produced so far only by methods of direct gene transfer[3,4], which require competent protoplasts[5]. Dividing cereal protoplasts, from which plants can be regenerated, are isolated from embryogenic cell cultures[6]. Such cultures are normally derived from immature embryos. We have established an embryogenic suspension culture from immature pollen grains[7] of Indica type rice *Oryza sativa* var. Chinsurah Boro II. Plants can subsequently be regenerated from protoplasts of these cultures[8]. We report here the transformation of Indica-type rice and transmission of the transgene to the progeny.

RESULTS

Microspore-derived embryogenic cell suspension culture. An embryogenic suspension culture was established via anther float culture from immature microspores of Indica-type rice as described recently[8] and maintained under diffuse light for isolation of protoplasts (Fig. 1a). The culture was composed of clusters of small, cytoplasm-rich cells and required subculturing every 5 days (Fig. 1b).

Protoplast culture, transformation and regeneration. Protoplasts were isolated from 5–12 month old cultures by incubation in an enzyme mixture that yielded an average 4×10^6 protoplasts/g cell suspension culture (Fig. 1c). No undigested clumps of cells were detected in the protoplast preparations. However, a few spontaneously fused protoplasts were observed. Nurse cultures or feeder layers were not required for culturing of isolated protoplasts, divisions and regeneration of plants (Fig. 1 d–h). PEG-mediated transformation was performed as described in the Experimental Protocol.

Treatment of protoplasts with DNA and PEG 6000 for a short time (10 min) was important in obtaining high frequencies of protoplast survival, divisions and subsequent plant regeneration. To allow selection for transformed cell clones, plasmid pGL2, which carries the *hph* gene under control of the 35S promoter and polyadenylation signal of CaMV, was used (Fig. 2a). The *hph* gene confers hygromycin resistance to transformed cells. Hygromycin resistant (Hm^r) cell clones were transferred to semi-solid N6 medium[9] containing 25 μg/ml hygromycin B, 1 mg/l 2,4-D and 0.3% agarose to allow for continued proliferation. No clones developed in the control samples (Fig. 1e, left dish). After 2–4 weeks, compact embryogenic clones were transferred to modified MS culture medium[10] without hygromycin B and somatic embryos developed (Fig. 1f). All cultures to this point were kept in the dark. Transfer of somatic embryos to hormone-free, modified MS medium under light led to the outgrowth of multiple shoots and roots (Fig. 1g). Plants regenerated from hygromycin-resistant clones grew to maturity in the greenhouse and set seeds (Fig 1h). These plants resembled control plants regenerated from untreated protoplasts or grown from seeds. Of the 77 plantlets recovered 24 were grown to maturity (Table 1).

Molecular data: Southern analysis. Total genomic DNA was isolated from 10 hygromycin-resistant primary regenerants and analysed by the method of Southern[11]. In all cases, hybridization with a radioactively-labeled probe specific for the protein coding region of *hph* demonstrated that this gene had integrated into the genome (data not shown). Figure 2b shows the Southern blot data for one representative primary transgenic plant (lanes 3–

TABLE 1 Recovery of transgenic Indica-rice plants from protoplasts treated with plasmid pGL2 and PEG.

Expt. No.	Antibiotic	Con. (μg/ml)	Plasmid	No. of Protoplasts Used ($\times 10^6$)	No. of Hm^r Clones (28d)	No. of Colonies Tested for Regeneration	Plants Regenerated (Green)	Plants Regenerated (Albino)
1	Hm	25	pGL2	6.5	200	60	—	—
2	Hm	25	pGL2	8.0	350	280	73	16
3	Hm	25	pGL2	8.5	116	22	4	—
4	Hm	25	pGL2	5.8	28	nt	nt	nt
C_1	—	—	—	8.0	2260*	430	25	2
C_2	Hm	25	—	8.0	—	—	—	—

*No. of colonies without selection; Hm=hygromycin; —indicates no response; nt=not tested; C_1 control without DNA and selection; C_2 control without DNA but with selection

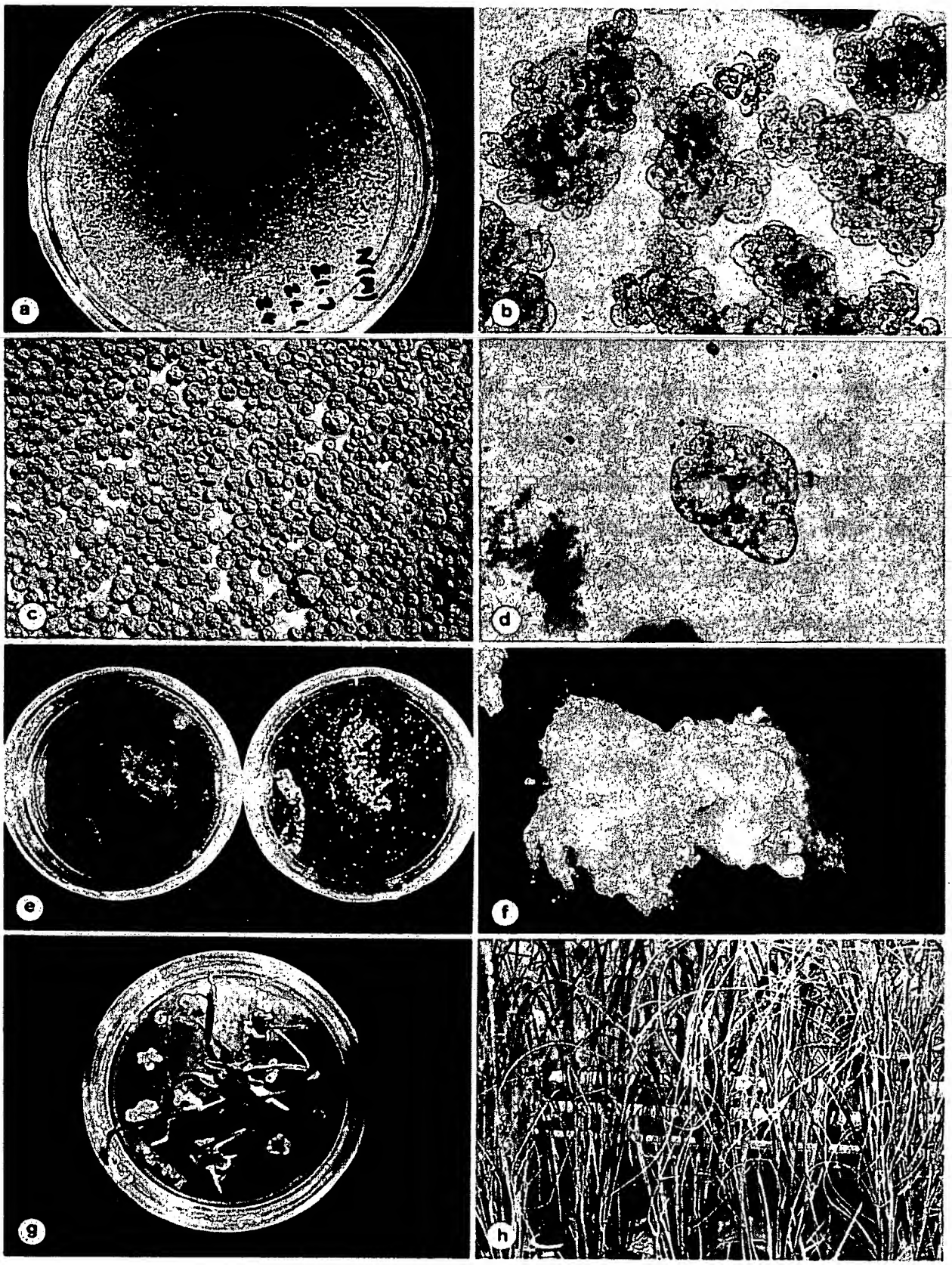


FIGURE 1 Regeneration of transgenic fertile indica-rice plants from protoplasts. (a) Embryogenic cell suspension (ECS) culture of microspore origin; (b) Enlarged view of ECS showing compact cytoplasmic-rich cell groups; (c) Protoplasts isolated from 5 month old ECS; (d) A protoplast derived group of cytoplasm-rich cells; (e) Hygromycin-resistant clones proliferating only from protoplast populations treated with plasmid pGL2 and PEG (right dish) and not in control treatment (left dish); (f) Resistant clones developing somatic embryos; (g) Differentiation of multiple shoots with roots in the light; (h) One primary transgenic plant with seeds and several offspring thereof.

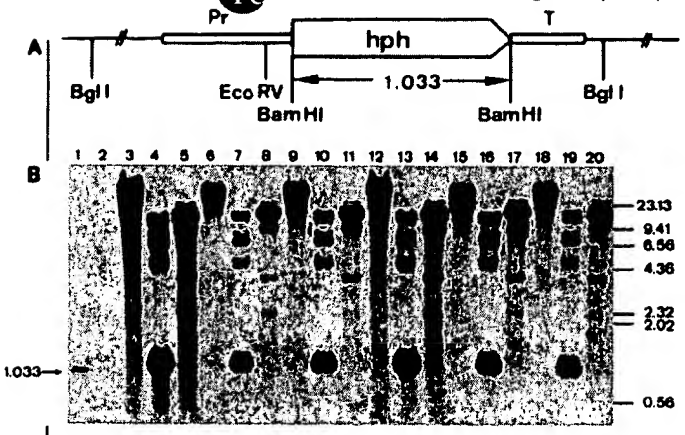


FIGURE 2 Molecular analysis of transgenic plants and offspring. (a) Part of plasmid pGL2 containing the bacterial gene for hygromycinphosphotransferase (*hph*) under the control of the expression signals of the 35S transcript of CaMV. For the transformation experiments plasmid pGL2 was cut with BglI releasing the fragment shown. The fragment contains a single EcoRV site within the promoter region; (b) Southern data are presented for one representative primary transgenic plant (lanes 3–5), and five offspring of this plant (lanes 6–8, 9–11, 12–14, 15–17, 18–20). Lane 1 represents a 3 copy reconstruction of the transforming plasmid cut with BamHI, lane 2 contains DNA from control (untransformed) rice plant, also cut with BamHI, samples are arranged in groups of three each. The first lane of these triplets (3, 6, 9, 12, 15, 18) contains undigested DNA, the second lane (4, 7, 10, 13, 16, 19) contains DNA restricted with BamHI (releasing a 1.033 kb fragment characteristic of the protein-coding region of the gene), the third lane (5, 8, 11, 14, 17, 20) contains DNA restricted with EcoRV (yielding border fragments between transforming DNA and host DNA). The size markers in kilo base pairs are derived from lambda DNA digested with HindIII.

5), and 5 independent offspring thereof (lanes 6–20). The integration pattern of the transgene in the offspring is identical to that of the parental plant. In all plants, hybridization of undigested DNA with the *hph*-specific probe showed that the transgene had integrated into the genome (Fig. 2b, lanes 3, 6, 9, 12, 15, 18). After restriction digestion with BamHI the expected 1.033 kb fragment characteristic of the coding sequence of *hph* was observed (Fig. 2b, lanes 4, 7, 10, 13, 16, 19; compare Fig. 2a). Digestion with EcoRV, which cuts once within the plasmid pGL2, produced junction fragments within the transforming DNA or with the host genome (Fig. 2b, lanes 5, 8, 11, 14, 17, 20). There was no hybridization to control material (Fig. 2b, lane 2).

Enzyme assay. Evidence demonstrating the presence of the transforming DNA and resistance to hygromycin B were complemented by enzymatic proof that the gene is functional; specific phosphorylation of the antibiotic was observed using protein extracts prepared from transgenic plants (Fig. 3). The plants taken for the enzyme assay were the same as those used for the Southern analysis shown in Figure 2b. The assay for hygromycin phosphotransferase activity was positive for the primary transgenic regenerant (Fig. 3, lane 1) and the five offspring plants analyzed (Fig. 3, lanes 3, 5, 7, 9, 11). The enzymatic evidence was also confirmed by growth of (selfed) seedlings derived from transgenic plants on 40 μg/ml hygromycin. They all developed into healthy plants under these conditions. So far, 31 seeds derived from 5 primary transgenic plants tested, were all hygromycin resistant, whereas 42 seeds from protoplast derived control plants were sensitive. Control plants germinated poorly, developed brown roots, ceased growth and died.

DISCUSSION

Transformation of rice, by electroporation[12,14,15] and by PEG[13] has been so far reported only with japonica types[12–15]. Electroporation has also been used for transformation in maize[16,17]. In these publications, authors report the absolute requirement of either nurse culture[15] or feeder layer cultures[16–18] for proliferation of rice and maize protoplasts. However, it is also possible that rice plants (both japonica and indica) could be obtained without any nurse or feeder cultures[8,13,19]. Moreover, nurse culture reduces the growth of the transformed colonies[14]. A heat-shock treatment (5 minutes at 45°C) to the recipient protoplasts prior to addition of the plasmid followed by PEG as reported to be beneficial for rice transformation[13,14] did not improve the transformation frequencies in our experiments with Indica-rice (data not shown). Perhaps, microspore-derived embryogenic cell suspension are, to our experience, well suited for reproducible production of transgenic colonies (Table 1). Transgenic plants obtained from protoplasts-derived clones via somatic embryogenesis resembled seed-derived plants.

Southern data are presented for one representative primary transgenic rice plant and five offspring of this plant. The data demonstrate (a) integration of the transforming plasmid DNA into high molecular weight DNA; (b) presence of the expected 1.033 kb BamHI fragment in the primary transgenic plant and offspring; and (c) hybrid fragments between rice DNA and transforming DNA. From the comparison of the intensity of the 1.033 kb BamHI fragment of the transgenic plants with the corresponding fragment of the control plasmid in a reconstitution experiment we estimate that approximately 50–100 copies of the plasmid are present in the transformants. Southern data obtained from five independent offspring plants derived from the same primary regenerant demonstrated identical integration pattern of the transforming DNA in all plants (Fig. 2b). Southern blot analysis of five additional offspring plants derived from the same primary transformant revealed exactly the same results (data not shown). All 10 progeny plants derived from the same primary regenerant were Hm^r. Two hypotheses could explain these data: either on the basis of a hemizygous primary transformant having integrated the transforming DNA into two or more independent genomic loci as described for tobacco[20]. In this case, a segregation of the hybridizing bands resolved on the Southern blots should be expected in the offspring plants, which was not observed (Fig. 2b). Alternatively, since the starting material for protoplast isolation and transformation was a microspore derived cell suspension and thus eventually being haploid at the time used for transformation, the recovered fertile primary transgenic plants could be homozygous. The lack of segregation of the integration pattern of the foreign DNA in the offspring indicates that the primary transformant is indeed homozygous. Therefore, information concerning the number of independent genomic integration sites can not be obtained from the analysis of first self-pollinated R_1 progeny. Further genetic analyses of more transgenic plants of independent experiments and determination of the ploidy levels of the cell suspension before transformation should clarify this point.

Resistance to hygromycin B is based on inactivation of the antibiotic via the transfer of the γ-phosphate from ATP[21]. The enzyme activity was absent in control rice tissue and present in both transformed leaves of the primary regenerant described here and in progeny plants thereof (Fig. 3). These data show that these plants not only carry the transgene but also express it.

We report here a simple and reproducible method of

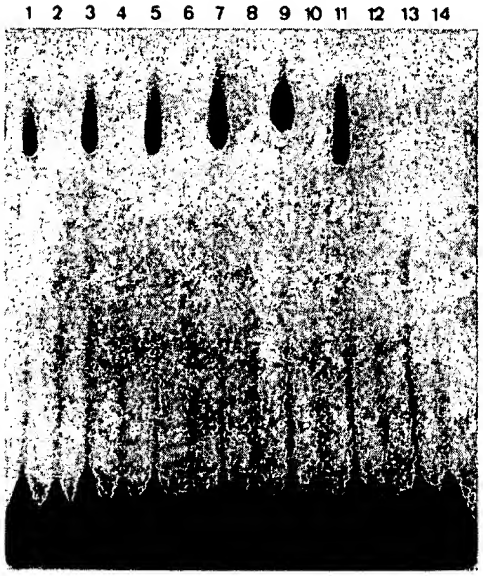


FIGURE 3 Enzyme assay for hygromycinphosphotransferase of primary transgenic and offspring plants. The plants assayed are identical to those used in the Southern blot analysis shown in Figure 2b. The autoradiogram of the TLC plate shows labelled ATP at the start in all lanes and labelled hygromycin B in lanes 1, 3, 5, 7, 9, 11. Even numbers represent enzyme reactions carried out without Hygromycin B, odd numbers represent reactions containing Hygromycin B. Lanes 1, 2 contain crude protein extract derived from the primary transgenic plant; lanes 3–12 contain extracts derived from five offspring of the same primary plant; lanes 13 and 14 contain extracts isolated from an untransformed control rice plant.

transformation of an important food crop, Indica-type rice and transmission of the foreign gene to the progeny. As co-transformation has been established as an efficient and routine procedure for the recovery of plants transgenic for non-selectable genes[15,22] and transgenic plants resistance to insect pests have already been tested and grown in the field[23], it should now be feasible to approach the production of Indica-type rice plants transgenic for agronomically more interesting genes.

EXPERIMENTAL PROTOCOL

Plant material. Experiments were carried out with one Indica-type rice variety (Chinsurah Boro II, IRRI Acc. No. 11484, kindly supplied by the International Rice Research Institute, Manila, Philippines and Rice Research Station at Chinsurah, West Bengal, India). Growth of donor plants, culture of microspores, embryogenic cell suspensions and protoplasts have been described earlier[7,8].

Plasmid construct. Plasmid pGL2 was obtained by cloning the *hph* gene as a BamHI fragment derived from plasmid pGL88[21] (under control of 35S promoter and polyadenylation signal of CaMV[24]) into the BamHI site of plasmid pDH51[25]. The transforming DNA used for direct gene transfer was cut with BglI releasing the fragment shown in Figure 2a. Plasmid pGL2 contains a single EcoRV site in the promoter region of the gene. Carrier DNA was prepared by dissolving calf thymus DNA (Sigma) sheared to an average size of approximately 4 kb in water and filter sterilized.

Protoplast culture, transformation and production of transgenic plants. An embryogenic cell suspension was obtained from dividing microspores and maintained for more than a year under diffuse light (7 $\mu E\ s^{-1}\ m^{-2}$ at 80 rpm) for 24 h photoperiod in a medium described earlier[8]. Protoplasts were isolated from 5-20 month old cultures by incubation in the following enzyme mixture: 4% cellulase onozuka RS (Kinki Yakult, Japan), 1% macerozyme R10 (Kinki Yakult, Japan), 0.02% pectolyase Y-23 (Seshin Pharmaceutical, Japan) (all w/v) dissolved in 0.4 M mannitol, 6.8 mM $CaCl_2$, pH 5.6 adjusted before filter sterilization. Protoplasts were washed three times by sedimentation in 0.4 M mannitol, 6.8 mM $CaCl_2$ and resuspended in a final density of 1.5×10^6/ml. Aliquots of 0.4 ml of the protoplast suspension were mixed with BglI-digested pGL2 plasmid DNA (6 µg) and calf thymus, carrier DNA (28 µg). Immediately after mixing the protoplasts with DNA, 0.4 ml of PEG solution (polyethyleneglycol MW 6000, Merck, 40% w/v) was added dropwise and the mixture of protoplasts, DNA and PEG incubated at 20°C for 10 min[26]. Protoplasts were then cultured with 8 ml of washing solution (0.4M mannitol, 0.1% MES, pH 5.6) slowly and centrifuged to remove the PEG. Subsequently, aliquots of 0.5 ml protoplasts suspension were mixed gently (in a 35 mm Falcon dish) with 0.5 ml of modified N6 medium containing 2.4% (w/v) agarose (Sea Plaque, FMC), 1 mg/l 2,4-D, 0.4 M sucrose, 28 mM glucose and allowed to gel. Following incubation in the dark at 27°C for 7 days, the agarose gel was cut into segments, which were transferred to 5 ml N6 medium with the following modifications: 1 mg/l, 2,4-D, 175 mM sucrose, 28 mM glucose and 0.3% (w/v) agarose and continued in 6 cm Sterilin plastic plates as bead type culture[27]. Hygromycin B was added to the medium at day 14 after protoplast isolation to a final concentration of 25 µg/ml. After 4–5 weeks of selection, visible colonies were transferred onto soft agarose N6 medium containing 1 mg/l 2,4-D, 175 mM sucrose, 28 mM glucose, 0.3% agarose, 25 µg/ml hygromycin B. Following 2–4 weeks at 24°C in the dark, cell colonies of ca. 1.5 mm diameter and developing somatic embryos were transferred to modified MS[10] culture medium containing 146 mM sucrose, 2 mg/l kinetin, 1 mg/l NAA, 0.8% agarose, 300 mg/l casein hydrolysate (ICN). Incubation in the light (24 $\mu E\ s^{-1}\ m^{-2}$), 16 h photoperiod led to the development of plantlets with multiple tillers with roots from somatic embryos. These were transferred to MS agar medium without hormones and containing 58 mM sucrose for 2–3 weeks. At this stage the agar was removed from the roots and the plantlets were transferred to potting compost and adjusted to greenhouse conditions. In the greenhouse, the plants grew to maturity and set seeds.

Progeny test for resistance to hygromycin. Seeds derived from self pollinated plants were surface sterilized in 1.8% (v/v) sodium hypochlorite and washed extensively with sterile distilled water. MS culture medium without hormones and vitamins either liquid or solidified with 0.8% agar, 1.0% sucrose, containing 40 µg/ml hygromycin B was used for germination of seeds. Seeds were incubated under light (24 $\mu E\ s^{-1}\ m^{-2}$), 16 h photoperiod at 24°C for 14 days.

Molecular analysis of transgenic plants. Total genomic DNA was isolated from leaf tissue of hygromycin-resistant and control plants. The leaves were freeze dried and ground in a mortar and pestle until a powder was obtained. DNA was extracted using the CTAB method[28]. Three µg of genomic DNAs was digested with BamHI or EcoRV restriction enzymes. Following electrophoresis through 0.8% agarose, DNA was transferred to Hybond-N nylon membranes. Hybridizations were done according to the instruction of the manufacturer (Amersham). The radioactive probe was prepared by the random primer method using (α-^{32}P)dATP[29]. The probe consisted of the protein-coding region of the *hph* gene (1.033 kbp BamHI fragment isolated from plasmid pGL2).

Assay for hygromycin phosphotransferase. The enzyme assay was carried out as previously described[30], modified according to personal communication by R. D. Shillito (CIBA Geigy, Research Triangle Park, North Carolina, USA). Leaves were frozen in liquid nitrogen and ground with pestle and mortar in extraction buffer (0.05 M Tris-HCl, pH 7.0, 10% glycerol, 0.1 mM phenylmethyl sulphonyl fluoride) (100–200 mg tissue/100 µl) in the presence of acid washed sea sand at 4°C. The samples were then centrifuged at 14000 rev/min for 5 min at 4°C and the supernatant used for the reaction. The enzyme reactions were carried out in 10 µl volumes containing 50 mM Tris-maleate, pH 7.0, 50 mM $CaCl_2$, 0.05 mM ATP, 0.4 µl (γ-^{32}P)ATP (10 mCi/ml; 3000 Ci/mmol), 62 µg hygromycin B and 5.6 µl crude extract. Reactions were carried out with and without hygromycin. Incubation was for 30 min at 37°C. One µl aliquots from the reaction mixtures were applied to a PEI-cellulose F TLC plate (Merck), which was developed in 50 mM sodium formate/formic acid, pH 5.4. The plates were dried prior to autoradiography.

Acknowledgments

We thank Paul Frick for his excellent care of the rice plants in the greenhouse and Laura Sarokin, J. Paszkowski and M. W. Saul for their helpful discussions and critical reading of the manuscript.

Received 15 April 1990; accepted 1 June 1990.

References

1. Khush, G. S. 1984. IRRI breeding program and its worldwide impact on increasing rice production, p. 61–94. *In:* Gene Manipulation in Plant Improvement. Gustafson, J. P. (Ed.). Plenum Press, NY.
2. Swaminathan, M. S. 1982. Biotechnology research and third world agriculture. Science **218**:967–972.
3. Potrykus, I. 1990. Gene transfer to cereals: An assessment. Bio/Technology **8**:535–542.
4. Paszkowski, J., Shillito, R. D., Saul, M. W., Mandak, V., Hohn, T., Hohn, B., and Potrykus, I. 1984. Direct gene transfer to plants. EMBO J **3**:2717–2722.
5. Potrykus, I. 1990. Gene transfer to plants: assessment and perspectives. Physiol. Plant. 73(1), *In press.*
6. Vasil, I. K. 1988. Progress in the regeneration and genetic manipulation of cereal crops. Bio/Technology **6**:397–402.
7. Datta, S. K., Datta, K., and Potrykus, I. 1990. Embryogenesis and plant regeneration from microspores of both "Indica" and "Japonica" rice *(Oryza sativa)*. Plant Sci. **67**:83–88.
8. Datta, S. K., Datta, K., and Potrykus, I. 1990. Fertile Indica rice plants from protoplasts of microspore derived suspension culture. Plant Cell Rep. *In press.*
9. Chu, C. C., Wang, C. C., Sun, S. S., Hsu, C., Yin, K. C., Chu, C. Y., and Bi, F. Y. 1975. Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. Sci. Sin. **18**:659–668.
10. Murashige, T. and Skoog, F. 1962. A revised medium for a rapid growth and bioassay with tobacco tissue culture. Physiol. Plant. **15**:473–479.
11. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
12. Toriyama, K., Arimoto, Y., Uchimiya, H., and Hinata, K. 1988. Transgenic rice plants after direct gene transfer into protoplasts. Bio/Technology **6**:1072–1074.
13. Zhang, W. and Wu, R. 1988. Efficient regeneration of transgenic rice plants from rice protoplasts and correctly regulated expression of foreign genes in the plants. Theor. Appl. Genet. **76**:835–840.
14. Zhang, H. M., Yang, H., Rech, E. L., Golds, T. J., Davis, A. S., Mulligan, B. J., Cocking, E. C., and Davey, M. R. 1988. Transgenic rice plants produced by electroporation-mediated plasmid uptake into protoplasts. Plant Cell Rep. **7**:379–384.
15. Shimamoto, K., Terada, R., Izawa, T., and Fujimoto, H. 1989. Fertile transgenic rice plants regenerated from transformed protoplasts. Nature **338**:274–277.
16. Fromm, M. E., Taylor, L. P., and Walbot, W. 1986. Stable transformation of maize after gene transfer by electroporation. Nature **319**:791–793.
17. Rhodes, C. A., Pierce, D. A., Mettler, D. M., Mascarenhas, D., and Detmer, J. J. 1988. Genetically transformed maize plants from protoplasts. Science **240**:204–207.
18. Lee, L., Schroll, R. E., Grimes, H. D., and Hodges, T. K. 1989. Plant regeneration from indica rice *(Oryza sativa)* protoplasts. Planta **178**:325–333.
19. Peterhans, A., Datta, S. K., Datta, K., Godali, G. J., Potrykus, I., and Paszkowski, E. 1990. Recognition efficiency of Dicotyledoneae specific promoter and RNA processing signals in rice. Mol. Gen. Genet. *In press.*
20. Potrykus, I., Paskowski, J., Saul, M. W., Petruska, J., and Shillito, R. D. 1985. Molecular and general genetics of a hybrid foreign gene introduced into tobacco by direct gene transfer. Mol. Gen. Genet. **199**:167–177.
21. Blochinger, K. and Diggelmann, H. 1984. Hygromycin B phosphotransferase as a selectable marker for DNA transfer experiments with higher eucaryotic cells. Mol. Cell. Biol. **4**:2929–2931.
22. Schocher, R. J., Shillito, R. D., Saul, M. W., Paskowski, J., and Potrykus, I. 1986. Co-transformation of unlinked foreign genes into plants by direct gene transfer. Bio/Technology **4**:1093–1096.
23. Delanny, X., LaVallee, B. J., Proksch, R. K. Fuchs, R. L., Sims, S. R., Greenplate, J. T., Marrone, P. G., Dodson, R. B., Augustine, J. J., Layton, J. G., and Fischhoff, D. A. 1989. Field performance of transgenic tomato plants expressing the *Bacillus thuringiensis* var. kurstaki insect control protein. Bio/Technology **7**:1265–1269.
24. Gronenborn, B. 1989. The molecular biology of cauliflower mosaic virus and its application as plant gene vector, p. 1–29. *In:* Plant DNA Infectious Agents. Hohn T. H. and Schell J. (Eds.). Springer-Verlag, Wien, NY.
25. Pietrzak, M., Shillito, R. D., Hohn, T., and Potrykus, I. 1986. Expression in plants of two bacterial antibiotic resistance genes after protoplast transformation with a new plant expression vector. Nucleic Acid Res. **14**:5857–5868.
26. Negrutiu, I., Shillito, R., Potrykus, I., Biasini, G. and Sala, F. 1987. Hybrid genes in the analysis of transformation condition. Plant Mol. Biol. **8**:363–373.
27. Shillito, R. D., Paskowski, J., and Potrykus, I. 1983. Agarose plating and a bead type culture technique enable and stimulate development of protoplasts-derived colonies in a number of plant species. Plant Cell Rep. **2**:244–247.
28. Murray, H. G. and Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acid Res. **8**:4321–4326.
29. Feinberg, A. P. and Vogelstein, B. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
30. Cabanes-Bastos, E., Day, A. G., and Lichtenstein, C. P. 1989. A sensitive and simple assay for neomycin-phosphotransferase II activity in transgenic tissue. Gene **77**:169–177.

Transgenic Research (1996) **5**: 49–55

Expression of a cystine-rich fish antifreeze in transgenic *Drosophila melanogaster*

BERNARD P. DUNCKER[1], J. ANNE HERMANS[2‡], PETER L. DAVIES[1,2] and VIRGINIA K. WALKER[1*]

Departments of Biology[1] and Biochemistry[2], Queen's University, Kingston, Ontario, K7L 3N6, Canada (Fax 1 613 545 6617)

Received 5 December 1994; revised 15 March 1995; accepted 16 March 1995

We have used *Drosophila melanogaster* as a model system for the transgenic expression of cystine-rich Type II antifreeze protein (AFP) from sea raven. This protein was synthesized and secreted into fly haemolymph where it migrated as a larger species (16 kDa) than the mature form of the protein (14 kDa) as judged by immunoblotting. *Drosophila*-produced Type II AFP demonstrated antifreeze activity both in terms of thermal hysteresis (0.13 °C) and inhibition of ice recrystallization. Recombinant AFP was purified and N-terminal sequencing revealed a 17 aa extension that began at the predicted signal peptide cleavage point. The expression of all three AFP types in transgenic *Drosophila* has now been achieved. We conclude that the globular Type II and Type III AFPs are better choices for antifreeze transfer to other organisms than is the more widely used linear Type I AFP.

Keywords: proprotein; secretion; haemolymph; thermal hysteresis

Introduction

Organisms that inhabit freezing environments have developed a number of approaches to protect their body fluids from unrestrained ice formation (Storey and Storey, 1988). One such strategy is to synthesize antifreeze proteins (AFPs). Although structurally diverse, AFPs from different species act in a similar fashion (Davies and Hew, 1990). They all inhibit freezing by binding to nascent ice crystals, thereby making the addition of water to the ice lattice less favourable (DeVries, 1983). This results in a non-colligative lowering of the fluid freezing point below the melting point, which is known as thermal hysteresis. AFPs also inhibit ice recrystallization (Knight, *et al.*, 1984). This is important for freeze-tolerant organisms because it reduces the size of ice crystals that form upon thawing, and thus limits the damage done to surrounding tissues (Knight and Duman, 1986).

The transgenic expression of AFPs offers promise as a means of conferring freeze resistance to species that do not normally synthesize proteins with thermal hysteresis activity. To date, antifreeze genes have only been isolated from fish, but some of these genes, or their synthetic variants, have already been used in transgenic applications. AFP-producing transgenic tobacco lines have been generated in the hopes of conferring frost resistance (Hightower *et al.*, 1991; Kenward *et al.*, 1993), and AFP genes have also been transferred to salmon with the aim of increasing their survival in ice-laden seawater (Shears *et al.*, 1991). While the AFP expression in these commercially important species is encouraging, measurable resistance to freezing has yet to be achieved. We have adopted the fruit fly *Drosophila melanogaster* as a model system to evaluate the transgenic expression of fish AFP genes. The alanine-rich Type I AFP of winter flounder was the first type to be produced in flies (Rancourt *et al.*, 1987). Although synthesis of this protein was detected on immunoblots, AFP levels were insufficient to register thermal hysteresis activity in the haemolymph. Subsequently, this protein has been produced in transgenic tobacco where levels of accumulation have been equally disappointing (Kenward *et al.*, 1993). The failure of the protein to accumulate has been attributed in part to its lack of tertiary structure and a resultant susceptibility to proteolysis.

*To whom correspondence should be addressed.

‡Present address: Agriculture Canada, Siebens-Drake Research Institute, 1400 Western Road, London, ON, N6G 2V4, Canada

Far greater success was achieved when the Type III AFP gene from Atlantic wolffish was expressed in *Drosophila* (Rancourt *et al.*, 1990). This gene codes for a tightly structured, globular protein that has no particular amino acid bias (Chao *et al.*, 1993; Sönnichsen *et al.*, 1993). Levels of circulating AFP in excess of 1 mg ml^{-1} were obtained and resulted in haemolymph thermal hysteresis values of up to 0.35 °C.

We have investigated a third candidate for transgenic expression, the Type II AFP gene of sea raven. Its antifreeze is characteristically rich in cysteinyl residues, all of which form disulfide bridges in the folded protein (Ng and Hew, 1992). In fish, Type II AFP is initially synthesized as a 163 aa protein, but the circulating form has 129 aa residues (Hayes *et al.*, 1989). Current signal peptide cleavage algorithms (von Heijne, 1986) predict an intermediate of 146 aa, and a protein of this length is produced in fall armyworm cells transformed with recombinant baculovirus containing Type II AFP cDNA (Duncker *et al.*, 1994). In this report we examine the synthesis, processing, and activity of Type II AFP in the whole animal model system and compare its suitability for transgenic applications to that of the other two AFP types.

Materials and methods

AFP expression construct P[w, YP1:srAFP]

A 550 bp *Hin*dIII-*Pst*I fragment (Fig. 1b) derived from sea raven AFP cDNA clone C2-1 (Ng *et al.*, 1986) was ligated into pTZ19R (Pharmacia). To facilitate a subsequent *Bgl*I/*Bgl*I ligation step, the *Bgl*I sequence at position +133 relative to the cap site was altered by site-directed mutagenesis from GCCATGATGGC to an alternate *Bgl*I sequence, GCCACCTTGGC. A fusion was then made between the mutagenized sea raven AFP cDNA coding for exons 1–4 and the genomic AFP DNA coding for intron 4 and exon 5 (Fig. 1c). This was achieved by the three-way directional ligation of (i) the 550 bp *Hin*dIII/*Pst*I cDNA fragment; (ii) the 640 bp *Eco*RI/*Pst*I fragment (Fig. 1a) of the 2.4 kb *Eco*RI/*Hin*dIII subclone of genomic clone SR7 (Hayes *et al.*, 1989); into (iii) *Eco*RI/*Hin*dIII-cut pUC19 to give psrAFP C/G. The 1639 bp *Bam*HI/*Pst*I fragment from pDMYP1/2:IR[a] (Riddell *et al.*, 1981) was ligated into pUC9, generating pYP1,2 (BP) (not shown). The 370 bp *Bgl*I/*Nsi*I YP fragment from pYP1,2 (BP) and the 1 kb *Bgl*I/*Eco*RI AFP fragment from psrAFP C/G were directionally ligated into the *Eco*RI and *Pst*I sites of the pW8 *P*-element vector polylinker (Klemenz *et al.*, 1987), to generate P[w, YP1:srAFP] (Fig. 1d). Although they have compatible cohesive ends for ligation, neither the *Pst*I site nor the *Nsi*I site is regenerated. This final ligation was facilitated through the earlier mutagenesis of the AFP *Bgl*I site to make the two *Bgl*I termini compatible.

Embryo microinjections and fly husbandry

Pre-syncytial blastoderm embryos collected from the flightless, white-eyed *Drosophila* host strain yw^{67};Ifm(3)3 (Duncker *et al.*, 1993) were microinjected with a mixture of P[w, YP1:srAFP] and transposase expression vector pπ25.7 wc (Spradling, 1986) in injection buffer (5 mM KCl, 0.1 mM NaH_2PO_4, pH 6.8), at concentrations of 350 ng λl^{-1} and 50 ng λl^{-1}, respectively, following standard procedures (Rubin and Spradling, 1982). Surviving G_0 flies were mated with injection stock flies, and the red eye colour of resulting progeny was used to identify transgenic G_1 flies. Homozygous lines were obtained by monitoring the eye phenotype through single pair matings.

Collection of fly haemolymph

Newly emerged (<24 h) adult flies were collected and placed on fresh food (10% yeast–sucrose medium) (Walker, 1985) for 48 h to ensure maximal *yp1* promoter-directed expression (Bownes *et al.*, 1988). For small-scale isolations of haemolymph, 50–100 flies were placed in 1.5 ml microcentrifuge tubes and frozen in dry ice for 5–10 min. For large-scale isolations, 1000–5000 flies were frozen in 50 ml centrifuge tubes. Following vigorous shaking to detach heads and legs, the flies were transferred to either 1 ml pipette tips (small-scale) or 5 ml syringes (large-scale) plugged with glass wool and were centrifuged (10 min, 10 000 × g, 4 °C) to collect the extruded haemolymph. PMSF was added to all samples at a final concentration of 5 mM. Haemocytes were pelleted by a second round of centrifugation (5 min, 14 000 × g, 4 °C) and the supernate was transferred to a fresh tube which was stored at −20 °C until needed.

Antifreeze activities

Ice crystal morphology and thermal hysteresis measurements were conducted using a nanolitre osmometer (Clifton Technical Physics, Hartford, New York, USA) following standard procedures (Chakrabartty and Hew, 1991). Photographs were taken with a WILD MPS 12 microcamera attachment through a Leitz Dialux 22 stereomicroscope, (Wild Heerbrugg, Heerbrugg, Switzerland). Thermal hysteresis is defined as the difference between melting and non-equilibrium freezing temperatures in °C. Ice recrystallization was observed by cryomicroscopy (Carpenter and Hansen, 1992).

Protein purification

Haemolymph (9 ml) collected from flies (50 g) of transgenic line P[w, YP1:srAFP]1 was applied to a Sephadex G-75 (Pharmacia) column (100 cm × 2.6 cm) equilibrated with 5 mM Tris-HCl, (pH 9.0). Aliquots of fractions were lyophilized, resuspended in 0.1 M

NH_4HCO_3 and assayed for thermal hysteresis activity. Pooled, active fractions were chromatographed on a Mono-Q FPLC column (Pharmacia) using a NaCl gradient in 5 mM Tris-HCl, (pH 9.0), at a flow rate of 0.5 ml min^{-1}. The NaCl gradient increased linearly in increments from 0 to 0.1 M in 10 min, from 0.1 M to 0.25 M in 30 min, and from 0.25 M to 0.5 M in 10 min. Fractions (0.5 ml) corresponding to the major Mono-Q peak were pooled and rechromatographed by reversed-phase HPLC on a C18 column (Vidac), using an acetonitrile gradient in 0.1% trifluoroacetic acid, at a flow rate of 1 ml min^{-1}.

Protein sequencing

Automated Edman degradations were done using an Applied Biosystems pulsed liquid sequencer (model 473A) equipped with microgradient phenylthiohydantoin analysis.

Results

Type II AFP is produced as a proprotein in transgenic Drosophila

The chimeric gene, P[w, YP1:srAFP], was constructed to place a sea raven AFP cDNA/genomic hybrid under the transcriptional control of the female-specific *Drosophila* *yp1* promoter (Fig. 1). Genomic sequence was incorporated to include an intron in the transcription unit because we have found that the presence of an intervening sequence boosts levels of Type III AFP transcripts in transgenic *Drosophila* (unpublished results). Two transgenic fly lines, P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2, were generated through embryo microinjections with P[w, YP1:srAFP]. Homozygous stocks were established and the integrity of their transgenes was confirmed by Southern analysis (Fig. 2). Three hybridization signals were observed on DNA blots of both transgenic lines. A common fragment at 2.3 kb, which was also present in the injection stock, represented endogenous *yp1*. A second common fragment, at 0.9 kb, confirmed the integrity of the internal *Pvu* II fragment of the insert. Both transgenic lines displayed an additional unique fragment, the size of which was determined by the location of the nearest genomic *Pvu* II site downstream of the different integration sites of the transferred vector.

Haemolymph collected from P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2 flies was subjected to SDS-PAGE and transblotted onto a nylon membrance (Millipore Immobilon-P). Immunoblots using rabbit anti-sea raven AFP antiserum (Ng *et al.*, 1986) revealed a cross-reacting 16 kDa protein in the haemolymph from female flies of both transgenic lines (Fig. 3, lanes b and d). No such

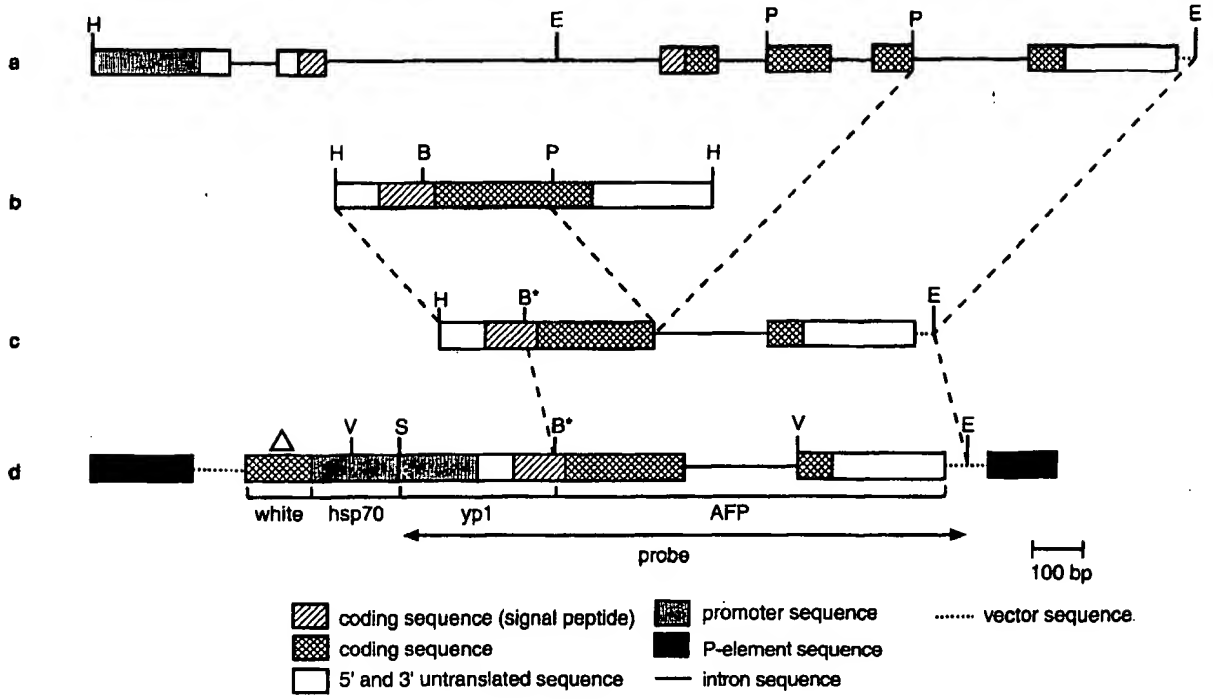


Fig. 1. Construction of sea raven AFP expression vectors. (a) sea raven AFP gene; (b) sea raven AFP cDNA; (c) cDNA/genomic fusion clone psrAFP C/G; (d) P-element expression construct P[w, YP1:srAFP]. Genes and constructs are drawn to scale, and pertinent restriction sites *Bgl* I (B), *Eco* RI (E), *Hin* dIII (H), *Pst* I (P), *Pvu* II (V) and *Sph* I (S) are indicated. The mutagenized *Bgl* I site is denoted by an asterisk. The double-headed arrow indicates the fragment (*Sph* I/*Eco* RI) used as a probe for Southern analysis. The triangle signifies that the white gene is not drawn to scale. Further details of the construction are presented in Materials and methods.

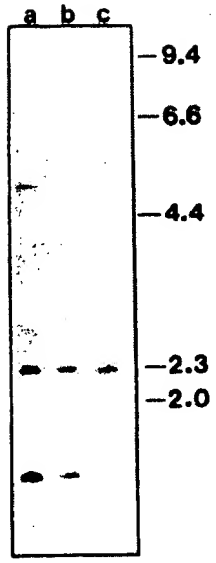


Fig. 2. Southern analysis of transgenic fly lines. Genomic DNA (10–20 λg) was isolated from flies of each line (Chia *et al.*, 1985) and digested with *Pvu* II. DNA fragments were separated by agarose gel electrophoresis and transferred to a nylon membrane. The fragment indicated in Fig. 1 was nick-translated and used as a probe. Lanes (a) and (b), DNA from transgenic lines, P[w, YP1:srAFP]1 and P[w, YP1:srAFP]2, respectively. Lane (c), DNA from the injection stock yw^{67};Ifm(3)3.

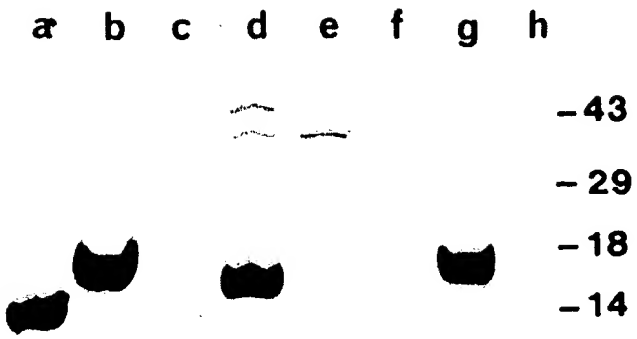


Fig. 3. Immunoblot analysis of insect-produced AFP. SDS-PAGE was performed using a 15% polyacrylamide SDS gel, containing 0.1 M sodium phosphate (pH 6.8) and 4 M urea. Prior to loading, all samples were boiled in loading buffer (0.1% SDS, 0.1 M sodium phosphate, 1% β-mercaptoethanol, 8 M urea, 0.003% bromophenol blue) for 5 min. Transfer to a nylon membrance was carried out. The membrane was incubated with rabbit anti-sea raven AFP antiserum and then with horseradish peroxidase-linked goat anti-rabbit IgG. Detection was performed using enhanced chemiluminescence (Amersham). Samples: purified AFP from sea raven serum (a); haemolymph (0.1 λl) from P[w, YP1:srAFP]1 females (b) and males (c); P[w, YP1:srAFP]2 females (d) and males (e); yw^{67};Ifm(3)3 injection stock females (f); culture medium (15 λl) from recombinant (g) and wild-type (h) baculovirus-infected fall armyworm cells. The position and size (kDa) of the protein standards are indicated at the side. The slight mobility difference between the observed AFPs in lanes b and d was an electrophoretic artifact of this particular gel. Additional high M_r bands are commonly, but not consistently seen for haemolymph samples challenged with this antiserum (compare lanes d and e to lanes b,c and f). It is not known what these bands represent.

protein was evident in the haemolymph of male transgenics (lanes c and e) or in that of nontransgenic, injection stock females (lane f). Mature Type II AFP purified from sea raven serum migrated as a 14 kDa protein (lane a). Culture medium from AFP-producing recombinant baculovirus-infected fall armyworm cells (lane g) as well as medium from wild-type baculovirus-infected cells (lane h) (Duncker *et al.*, 1994) were used as a standard and control, respectively. The baculovirus- and *Drosophila*-produced Type II AFPs demonstrated a similar mobility.

Recombinant proAFP demonstrates antifreeze activity

Upon cooling, ice crystals in the haemolymph samples from female P[w, YP1:srAFP]1 transgenics grew into hexagonal bipyramids (Fig. 4b). These were stable in size and, as noted by the slight curvature at the junction of the two hexagonal pyramids, characteristic of those seen in Type II AFP solutions from fish (Fig. 4a). In contrast, those observed in haemolymph samples from host nontransgenic female flies (yw^{67};Ifm(3)3) lacked this bipyramidal shape and demonstrated unrestrained growth (Fig. 4c). The thermal hysteresis value for female P[w, YP1:srAFP]1 haemolymph was determined to be 0.13 ± 0.2 °C.

These same haemolymph samples were evaluated for their ability to inhibit ice recrystallization. After snap

Fig. 4. Ice crystal morphology. Samples analysed were sea raven AFP purified from fish serum (a), haemolymph from female P[w, YP1:srAFP]1 (b) and female yw^{67};Ifm(3)3 (c) flies. Magnification was 700-fold.

freezing to −133 °C, the ice crystal size at a number of temperatures approaching the haemolymph melting point was recorded. When compared to haemolymph from nontransgenic flies, P[w, YP1:srAFP]1 haemolymph

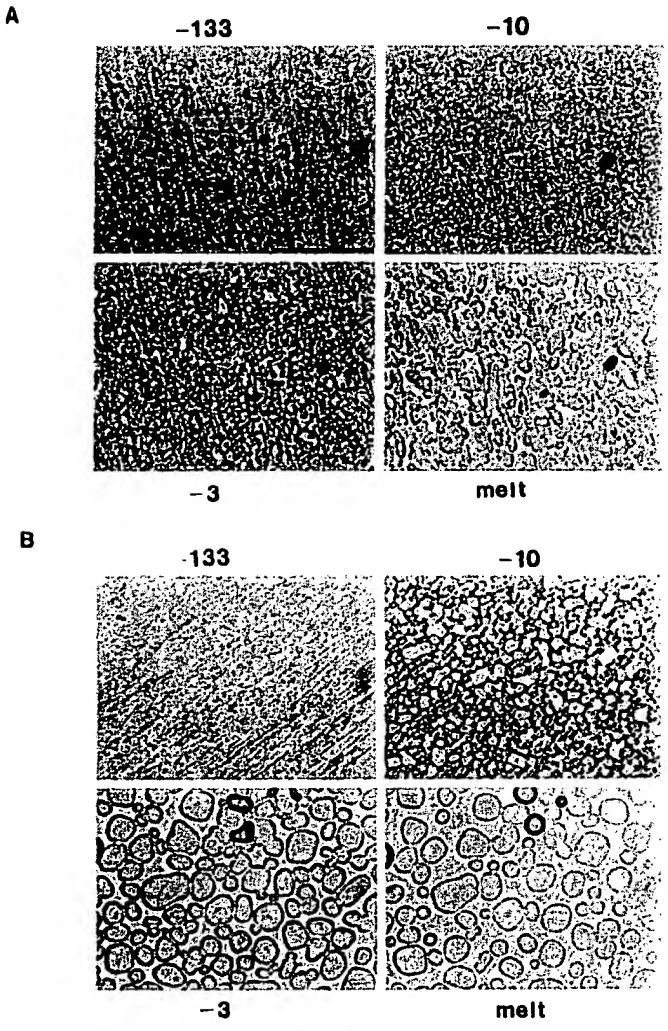


Fig. 5. Ice recrystallization assays. Haemolymph from female P[w, YP1:srAFP]1 (panel A) and yw^{67};Ifm(3)3 (panel B) flies was snap frozen at −133 °C and micrographs were taken at −133 °C, −10 °C, −3 °C, and at the onset of melting. The bar represents 50 λm; all micrographs are to the same magnification.

formed smaller ice crystals at −10 °C, −3 °C and at the onset of melting (Fig. 5), indicating inhibition of ice recrystallization.

Purification and characterization of Drosophila-produced Type II proAFP

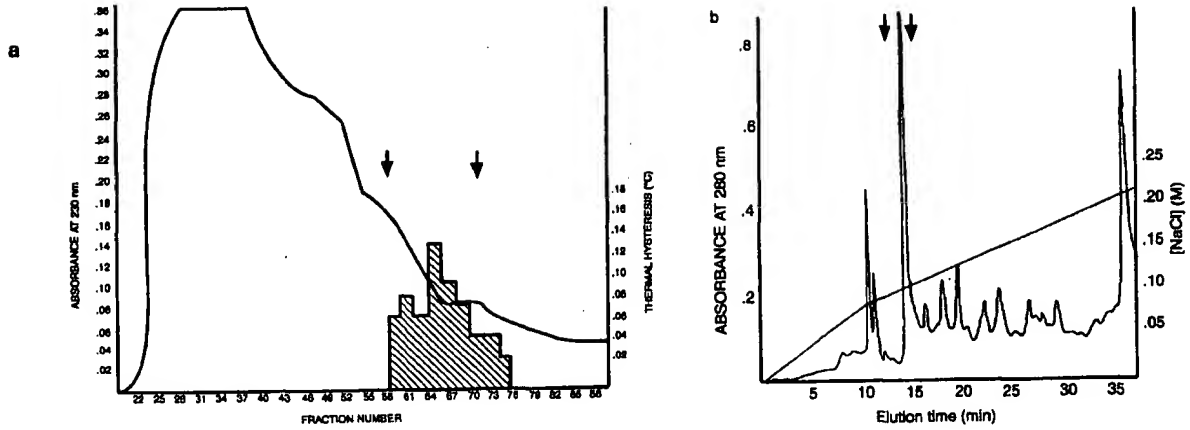
AFP from P[w, YP1:srAFP]1 haemolymph was purified by Sephadex G-75 gel permeation chromatography (Fig. 6a) followed by Mono-Q FPLC (Fig. 6b) and reversed-phase HPLC (Fig. 6c). The AFP was detected by thermal hysteresis activity in the low M_r end of the G-75 eluate, consistent with a M_r of 16000. On the Mono-Q column the activity co-chromatographed with the major A_{280}-absorbing peak but based on the subsequent HPLC profile was only a very minor constituent of it. Purified recombinant sea raven AFP was subjected to 10 cycles of automated Edman degradation and gave a N-terminal sequence of NDDKILKGTA. This matches the expected N-terminus for the proform of the recombinant AFP beginning at residue 18 of the 163 aa primary translation product, as predicted by the von Heijne rules for signal peptide cleavage (von Heijne, 1986).

Discussion

We have produced active Type II AFP in transgenic *Drosophila melanogaster*. To ensure efficient secretion into the haemolymph, the fusion between *yp1* and AFP gene sequences was made in the regions encoding their signal peptides in such a way as to preserve a hydrophobic core within the signal sequence. The YP1 signal sequence was changed from MNPMRVLSLLACLAVAALAK to MNPMRVLSLLACLAVAALA↓**LTQA**. The fusion point is indicated by the arrow. Using the rules of von Heijne (1986), signal peptide cleavage is predicted to occur after the Ala, four residues C-terminal to the fusion point of the hybrid signal peptide (above), just as it does in the native fish signal peptide. This strategy for designing a chimaeric signal sequence had previously been used for the correct processing and secretion of Type III AFP in transgenic *Drosophila* (Rancourt *et al.*, 1990).

As predicted, a 146 aa proprotein was produced in the flies. Since there was no further processing of the proAFP, it is likely that the proteinase activity responsible for pro region removal in sea raven is absent in *Drosophila*. The antifreeze activity of the haemolymph from transgenic flies was demonstrated by both thermal hysteresis and inhibition of ice recrystallization. Haemolymph from control flies lacked these activities. Since Type II AFP loses its activity under reducing conditions (Slaughter *et al.*, 1981) it appears that it is correctly folded through disulfide bridge formation. These observations are consistent with our previous experience using insect cell culture where an active proprotein was also synthesized (Duncker *et al.*, 1994).

This now represents the third antifreeze type to be produced in transgenic flies. Previously, winter flounder (Type I) and wolffish (Type III) AFP genes were used in expression constructs where each was placed under the transcriptional control of *Drosophila* yolk protein promoters. Type I could not be detected in transgenic flies (Rancourt *et al.*, 1992), although this protein was transiently synthesized under heat shock conditions when its gene was linked to the *Drosophila hsp70* promoter (Rancourt *et al.*, 1987). In contrast, fly lines carrying the Type III gene synthesized AFP levels that gave freezing point depressions of 0.1–0.35 °C. The fly lines in this report, P[w, YP1:srAFP]1 and 2, produced haemolymph thermal hysteresis readings of up to 0.13 °C. Type II and



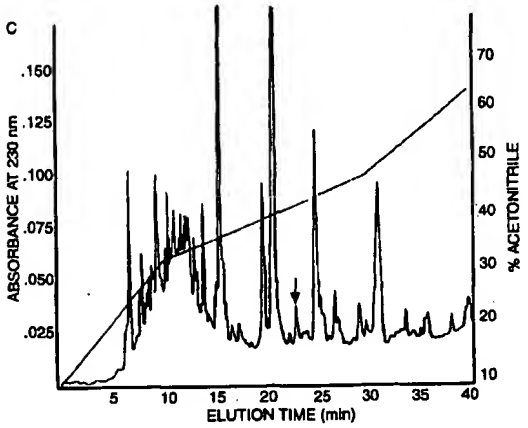


Fig. 6. Purification of *Drosophila*-produced AFP from line P[w, YP1:srAFP]1. (a) Sephadex G-75. Absorbance readings at 230 nm are indicated by the tracing while thermal hysteresis (freezing point depression) is represented by the histogram. Fractions between the arrows were pooled. (b) Ion exchange FPLC. Pooled G-75 fractions were chromatographed on a Mono-Q column (Pharmacia). Fractions between the arrows demonstrated antifreeze activity and were pooled (1.5 ml) for further purification. (c) Reversed-phase HPLC. Pooled FPLC fractions were loaded onto the HPLC column in solution A (0.1% trifluoroacetic acid) and eluted by a gradient of solution B (80% acetonitrile in 0.1% trifluoroacetic acid). The solution B concentration was increased linearly to 40% over 10 min, to 60% over a further 20 min, and to 80% over a subsequent 10 min. Fractions (1 ml) were collected. The greatest antifreeze activity was detected in the fraction corresponding to the peak indicated by the arrow (0.74 °C, for a tenfold concentration in 0.1 M NH_4HCO_3). This was the only peak showing significant thermal hysteresis. The large number of additional peaks likely represent oxidized haemolymph as well as melanin adducts and polymers from the insect wound response (Boman and Hultmark, 1987). Further details of purification steps are presented in Materials and methods.

III thermal hysteresis levels were therefore reasonably comparable, bearing in mind that the Type III expression construct included two copies of the AFP gene, whereas the Type II construct used in this study had only one. These results indicate that sea raven Type II AFP as well as the Type III AFP of wolffish are much better candidates in developing freeze resistance than the more widely used flounder Type I AFP (Cutler *et al.*, 1989; Georges *et al.*, 1990; Hightower *et al.*, 1991; Shears *et al.*, 1991; Kenward *et al.*, 1993). One possible reason for this difference is that Type I AFP has no tertiary structure and is less stable than the globular AFPs at room temperature; at −1 °C flounder AFP is 87% α-helical whereas at 25 °C the proportion drops to 47% (Ananthanarayanan and Hew, 1977). The lower α-helical content at room temperature might therefore contribute to a faster rate of degradation. In studies where a winter flounder Type I AFP gene was expressed in tobacco, AFP was detectable when the plants were grown at 4 °C, but not when they were placed at room temperature (Kenward *et al.*, 1993). In experiments with transgenic *Drosophila* expressing flounder AFP, we have found that this protein persists for much longer when flies are reared at 10 °C, rather than at room temperature (unpublished results, 1994).

The insights gained using this system have demonstrated its potential to assess candidate AFPs for use in other transgenic organisms. Additionally, it may have applications in the study of insect AFPs that, like the sea raven Type II AFP, appear to be extensively disulfide-bonded.

Acknowledgements

We would like to thank Dr Choy Hew (University of Toronto) for the gift of anti-sea raven AFP antisera, and

Dr John Carpenter (University of Colorado Health Science Centre) for cryomicroscopy. This research was supported by MRC and NSERC grants to P.L.D and V.K.W., respectively, and by an Ontario Graduate Scholarship awarded to B.P.D.

References

Ananthanarayanan, V.S. and Hew, C.L. (1977) Structural studies on the freezing point-depressing protein of the winter flounder *Pseudopleuronectes americanus*. *Biochem. Biophys. Res. Commun.* **74**, 685–9.

Boman, H.G. and Hultmark, D. (1987) Cell-free immunity in insects. *Ann. Rev. Microbiol.* **41**, 103–26.

Bownes, M., Scott, A. and Shirras, A. (1988) Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. *Development* **103**, 119–28.

Carpenter, J.F. and Hansen, T.N. (1992) Antifreeze protein modulates cell survival during cryopreservation: mediation through influence on ice crystal growth. *Proc. Natl Acad. Sci. USA* **9**, 8953–7.

Chakrabartty, A. and Hew, C.L. (1991) The effect of enhanced alpha-helicity on the activity of a winter flounder polypeptide. *Eur. J. Biochem.* **202**, 1057–63.

Chao, H., Davies, P.L. Sykes, B.D. and Sönnichsen, F.D. (1993) Use of proline mutants to help solve the NMR solution structure of type III antifreeze protein. *Protein Sci.* **2**, 1411–28.

Chia, W., Karp, R., McGill, S. and Ashburner, M. (1985) Molecular analysis of the *Adh* region of the genome of *Drosophila melanogaster*. *J. Mol. Biol.* **186**, 689–706.

Cutler, A.J., Saleem, M., Kendall, E., Gusta, L.V., Georges, F. and Fletcher, G.L. (1989) Winter flounder antifreeze protein improves the cold hardiness of plant tissues. *J. Plant Physiol* **135**, 351–4.

Davies, P.L. and Hew, C.L. (1990) Biochemistry of fish antifreeze proteins. *FASEB J.* **4**, 2460–8.

DeVries, A.L. (1983) Antifreeze peptides and glycopeptides in cold-water fishes. *Ann. Rev. Physiol* **45**, 245–60.

Duncker, B.P., Hermans, J.A., Davies, P.L. and Walker, V.K. (1993) Biological containment of transgenic fruit files. *Dros. Info. Serv.* **72**, 133.

Duncker, B.P., Gauthier, S.Y. and Davies, P.L. (1994) Cystine-rich fish antifreeze is produced as an active proprotein precursor in fall armyworm cells. *Biochem. Biophys. Res. Commun.* **203**, 1851–7.

Georges, G., Saleem, M. and Culter, A.J. (1990) Design and cloning of a synthetic gene for the flounder antifreeze protein and its expression in plant cells. *Gene* **91**, 159–65.

Hayes, P.H., Scott, G.K., Ng, N.F., Hew, C.L. and Davies, P.L. (1989) Cystine-rich type II antifreeze protein precursor is initiated from the third AUG codon of its mRNA. *J. Biol. Chem.* **264**, 18761–7.

Heijne, G. von (1986) A new method for predicting signal peptide sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683–91.

Hightower, R., Baden, C., Penzes, E., Lund, P. and Dunsmuir, P. (1991) Expression of antifreeze proteins in transgenic plants. *Plant Mol. Biol.* **17**, 1013–21.

Kenward, K.D., Altschuler, M., Hildebrand, D. and Davies, P.L. (1993) Accumulation of type I fish antifreeze protein in transgenic tobacco is cold-specific. *Plant Mol. Biol.* **23** 377–85.

Klemenz, R., Weber, U. and Gehring, W.J. (1987) The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucl. Acids Res.* **15**, 3947–59.

Knight, C.A. and Duman, J.G. (1986) Inhibition of recrystallization of ice by insect thermal hysteresis proteins: a possible cryoprotective role. *Cryobiology* **23**, 256–62.

Knight, C.A., Hallett, J. and De Vries, A.L. (1984) Fish antifreeze protein and the freezing and recrystallization of ice. *Nature* **308**, 295–6.

Ng, N.F.L. and Hew, C.L. (1992) Structure of an antifreeze polypeptide from the sea raven. *J. Biol. Chem.* **267**, 16069–75.

Ng, N.F.L., Khiet-Yan, T. and Hew, C.L. (1986) Structure of an antifreeze polypeptide precursor from the sea raven, *Hemitripterus americanus*. *J. Biol. Chem.* **261**, 15690–5.

Rancourt, D.E., Walker, V.K. and Davies, P.L. (1987) Flounder antifreeze protein synthesis under heat shock control in transgenic *Drosophila melanogaster*. *Mol. Cell. Biol.* **7**, 2188–95

Rancourt, D.E., Peters, I.D., Walker, V.K. and Davies, P.L. (1990) Wolffish antifreeze protein from transgenic *Drosophila*. *Bio/Technology* **8**, 453–7.

Rancourt, D.E., Davies, P.L. and Walker, V.K. (1992) Differential translatability of antifreeze protein mRNAs in a transgenic host. *Biochim. Biophys. Acta* **1129**, 188–94.

Riddell, D.C., Higgins, M.J., McMillan, B.J. and White, B.N. (1981) Structural analysis of three vitellogenin genes in *Drosophila melanogaster*. *Nucl. Acids Res.* **9**, 1323–38.

Rubin, G.M. and Spradling, A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–53.

Shears, M.A., Fletcher, G.L., Hew, C.L., Gauthier, S. and Davies, P.L. (1991) Transfer, expression and stable inheritance of antifreeze protein genes in atlantic salmon (*Salmo salar*). *Mol. Mar. Biol. Biotech* **1**, 58–63.

Slaughter, D., Fletcher, G.L., Ananthanarayanan, V.S. and Hew, C.L. (1981) Antifreeze proteins from the sea raven, *Hemitripterus americanus*. Further evidence for diversity among fish polypeptide antifreezes. *J. Biol. Chem.* **256**, 2022–6.

Sönnichsen, F.D., Sykes, B.D., Chao, H. and Davies, P.L. (1993) The nonhelical structure of antifreeze protein Type III. *Science* **259**, 1154–7.

Spradling, A.C. (1986) P-element mediated transformation. In: Roberts, D.B. (ed.) *Drosophila: A Practical Approach*. pp. 175–97 Oxford: IRL Press.

Storey, K.B. and Storey, J.M. (1988) Freeze tolerance in animals. *Physiol. Rev.* **68**, 27–84.

Walker, V.K. (1985) Yolk polypeptide gene expression in Minute *Drosophila* females. *Biochem. Genet.* **23**, 363–78.

ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY

Volume 362

ASPARTIC PROTEINASES

Structure, Function, Biology, and Biomedical Implications

Edited by Kenji Takahashi

Library of Congress Cataloging-in-Publication Data

Aspartic proteinases : structure, function, biology, and biomedical implications / edited by Kenji Takahashi.
p. cm. -- (Advances in experimental medicine and biology ; v. 362)
Proceedings of the 5t International Conference on Aspartic Proteinases, held Sept. 19-24, 1993 at the Naito Museum of Pharmaceutical Science and Industry, in Kawashima-cho, Gifu Prefecture, Japan.
Includes bibliographical references and index.
ISBN 0-306-44830-0
1. Aspartic proteinases--Congresses. I. Takahashi, Kenji. II. International Conference on Aspartic Proteinases (5th : 1993 : Naito Museum of Pharmaceutical Science and Industry) III. Series.
[DNLM: 1. Aspartic Proteinases--physiology--congresses. WI AD559 v. 362 1994 / QU 136 A8376 1994]
QP609.A86A88 1994
612'.01516--dc20
DNLM/DLC
for Library of Congress
94-43762
CIP

Proceedings of the Fifth International Conference on Aspartic Proteinases, held September 19–24, 1993, in Gifu, Japan

ISBN 0-306-44830-0

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

10 9 8 7 6 5 4 3 2 1

Printed in the United States of America

COMPARISON OF THE ACTIVE SITE SPECIFICITY OF THE ASPARTIC PROTEINASES BASED ON A SYSTEMATIC SERIES OF PEPTIDE SUBSTRATES

Ben M. Dunn, Paula E. Scarborough, W. Todd Lowther, and Chetana Rao-Naik

Department of Biochemistry and Molecular Biology
University of Florida College of Medicine
Gainesville, FL 32610-0245

INTRODUCTION

The first crystal structures of members of the aspartic proteinase family were discussed at the 1976 Conference in Norman, Oklahoma [1]. Since that time, an increasingly rapid stream of structural information has issued from the crystallography groups [2]. At this point it is clear that members of this family are closely related structurally, with a central core composed of two domains, each of which has a two-layered β-sheet arrangement (see Figure 1 for an overall representation of a typical aspartic proteinase).

At the interface of these two domains in all the enzymes is a clearly-defined active site cleft where substrate and inhibitor binding takes place. The dimensions of this cleft are longer (20-30 Angstroms) than wide, so that a peptide substrate binds in an extended fashion. From structural studies of bound inhibitor molecules at least seven amino acids can be accommodated within the cleft in an extended β-strand conformation. This agrees with early specificity studies [3,4] where demonstrable preferences for specific amino acids were observed in analyses of protein cleavage by porcine pepsin. Based on these reports, we designed and synthesized [5] a peptide substrate which proved of value in the kinetic analysis of enzymes in the aspartic proteinase class.

In addition, work initially described by Fruton [6] and more recently pursued by Hofmann and his colleagues [7,8] has clearly demonstrated that extending a peptide substrate through at least five amino acids leads to significant increases in k_{cat} when compared to shorter peptides. Thus, it is obvious that the extended cleft of the enzymes in the aspartic proteinase class has evolved to provide multiple points of interaction between substrate and enzyme in order to maximize efficiency of cleavage.

The increase in rate of cleavage is usually ascribed to added hydrogen bonds between enzyme residues and the backbone -NH- groups at P_3 and P_2' or P_3'. [Positions of a substrate or inhibitor are designated according to Schechter and Berger, BBRC (1967), 27:157]. Indeed,

Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications
Edited by Kenji Takahashi, Plenum Press, New York, 1995

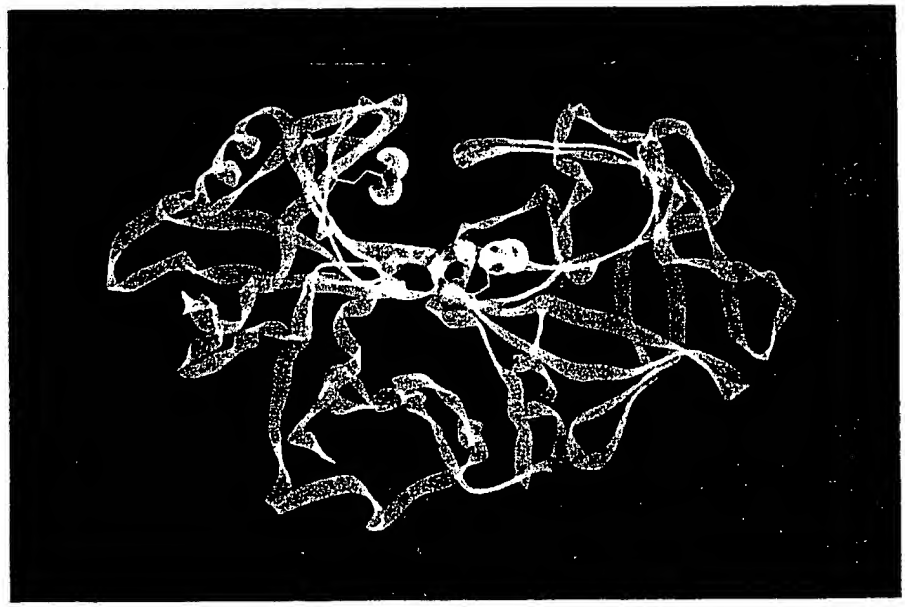


Figure 1. Ribbon diagram representation of the structure of an archetypal aspartic proteinase. The N-terminal domain is on the right in this view. The side chains of the two catalytic residues are represented by red stick structures, and the two oxygen atoms are indicated as van der Waals surfaces, seen as red dots. The positions of two important Glu residues in porcine pepsin (13 and 287) are also given by white stick structures, with the oxygen atoms of these indicated as white van der Waals surfaces. (See color plate preceding page 1.)

examination of all enzyme/inhibitor complexes has shown that almost all -NH- and C=O groups of a bound peptide or peptidomimetic are involved in hydrogen bonding. It is easy to imagine a picture where a bound peptide is "tied down" at multiple points by this hydrogen bonding. This would maximize residence time of a peptide within the active site cleft.

In recent years, we have turned our attention to a second question in these enzymes: How can we account for differences in specificity of cleavage between various members of this family? While it is clear that the extensive backbone hydrogen bonding mentioned above is one critical factor in the rate of cleavage, this cannot account for specificity, as all peptide sequences have identical -NH- and C=O groups.

To account for differences in the specific points of cleavage it is necessary to consider the second feature of a bound peptide or inhibitor: the side chain groups. While hydrogen bonding "ties down" a bound peptide, and may contribute to the energy of binding, the side chains of the peptide are able to interact with specific regions within the cleft, potentially with either a positive or a negative effect upon the energy of binding.

A schematic diagram illustrating the extended mode of binding of a peptide within the active site cleft of an aspartic proteinase as well as the concept that individual side chains of the bound peptide interact with defined regions ("subsites") of the cleft is presented in Figure 2.

We can then hypothesize that differences in the interactions at specific positions, when comparing two enzymes in this family interacting with one peptide substrate, can account for observed differences in the rates of cleavage of the peptide sequence. This hypothesis is based upon knowledge of the structures of these enzymes (and their complexes) which reveals that the overall features of the active site clefts are similar, while the specific

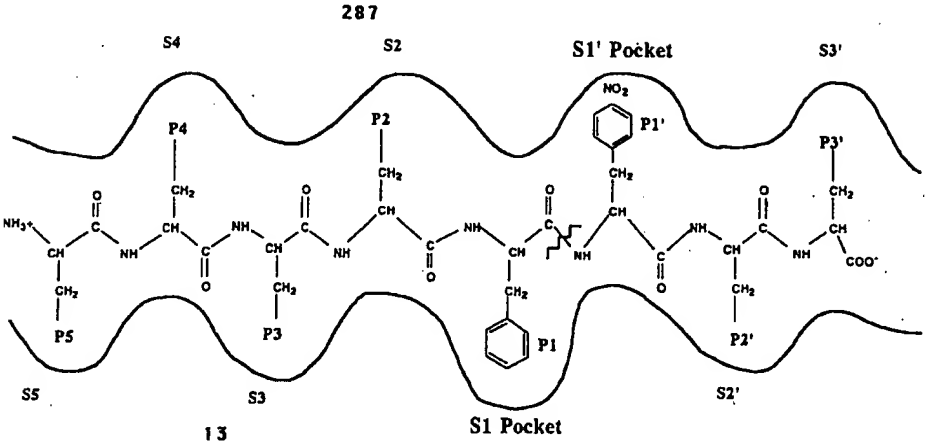


Figure 2. Schematic representation of the binding of a substrate in the active site cleft of an aspartic proteinase. Note the extended β-strand conformation of the substrate, in agreement with the pattern seen in inhibitor complexes. The scissile peptide bond is juxtaposed with the catalytic apparatus. Each amino acid side chain of the substrate interacts with a specific region of the active site, or "subsite". The involvement of residues 13 and 287 in the S_3 or S_2 regions, respectively, are indicated by those numbers in the appropriate

amino acid residues of the enzyme surface show some variation in positions that we believe to be critical to specificity.

In order to explore the differences in specificity between members of this family, and to test the hypothesis we have developed above, a set of peptide substrates was constructed. This set is based upon previous results with porcine pepsin [9], with the parent peptide having the sequence:

Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu

where the asterisk (*) indicates the point of cleavage, and Nph is the three-letter code for *p*-NO_2phenylalanine. The two aromatic residues, Phe in the P_1 position and Nph in the P_1' position, fit the primary specificity of most members of the aspartic proteinase family well. In the series given in Figure 3, we have kept the two primary specificity determining residues (-Phe*Nph-) constant, but have varied the residues in the other six positions to include the small, hydrophobic Ala, the small, hydrogen-bonding Ser, the small, acidic Asp, the larger basic residue Arg, and the larger, hydrophobic Leu. In this series, we have avoided the aromatic residues to avoid creating new cleavage points and the sulfur-containing amino acids in favor of long-term stability.

This report will compare selected results with porcine pepsin, human cathepsin D, and the aspartic proteinase from *Rhizopus chinensis*. The three-dimensional structures of those proteins are known at this point, facilitating interpretation of the differences observed.

MATERIALS AND METHODS

All peptides used in this study were synthesized by solid-phase methodology, utilizing *t*-BOC chemistry. Following cleavage from the synthetic resin by anhydrous HF, the purity of the peptides was assessed by amino acid analysis, HPLC (reversed-phase C-18

Peptide XVI Series

"Ben's Bookshelf"

LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
ALA- - - - - - - - * - - - - -	$XVIP_5A$
ASP- - - - - - - - * - - - - -	$XVIP_5D$
SER- - - - - - - - * - - - - -	$XVIP_5S$
LEU- - - - - - - - * - - - - -	$XVIP_5L$
ARG- - - - - - - - * - - - - -	$XVIP_5R$
LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
- -ALA- - - - - - * - - - - -	$XVIP_4A$
- -ASP- - - - - - * - - - - -	$XVIP_4D$
- -SER- - - - - - * - - - - -	$XVIP_4S$
- -LEU- - - - - - * - - - - -	$XVIP_4L$
- -ARG- - - - - - * - - - - -	$XVIP_4R$
LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
- - - -ASP- - - - * - - - - -	$XVIP_3D$
- - - -SER- - - - * - - - - -	$XVIP_3S$
- - - -LEU- - - - * - - - - -	$XVIP_3L$
- - - -ARG- - - - * - - - - -	$XVIP_3R$
LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
- - - - - -ALA- - * - - - - -	$XVIP_2A$
- - - - - -ASP- - * - - - - -	$XVIP_2D$
- - - - - -SER- - * - - - - -	$XVIP_2S$
- - - - - -LEU- - * - - - - -	$XVIP_2L$
- - - - - -ARG- - * - - - - -	$XVIP_2R$
LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
- - - - - - - - - * - -ALA- -	$XVIP_2'A$
- - - - - - - - - * - -ASP- -	$XVIP_2'D$
- - - - - - - - - * - -SER- -	$XVIP_2'S$
- - - - - - - - - * - -LEU- -	$XVIP_2'L$
LYS-PRO-ALA-LYS-PHE*NPH-ARG-LEU	XVI
- - - - - - - - - * - - - -ALA	$XVIP_3'A$
- - - - - - - - - * - - - -ASP	$XVIP_3'D$
- - - - - - - - - * - - - -SER	$XVIP_3'S$
- - - - - - - - - * - - - -ARG	$XVIP_3'R$

Figure 3. Comparison of the sequences of peptides created as an internally consistent set of substrates. Use of this set of peptides provides the opportunity to investigate the specificity of a newly discovered enzyme. Cleavage between Phe and Nph results in a chromogenic shift that can be quantitated.

with CH_3CN as the mobile phase), and, in some cases, capillary electrophoresis and mass spectroscopy. The typical preparation contained 90% of the desired peptide, with small amounts of peptides containing residual blocking groups as the most common contaminant.

The enzymes utilized herein have been described in previous publications from this laboratory[2].

Active site titration with the tightest binding inhibitor available for each enzyme was used to quantitate the concentration of enzyme in the assays.

In all cases, the assay depends on the shift in absorption due to cleavage of the -Phe*Nph peptide bond, and the resulting effect of the new $-NH_3+$ group on the *p*-nitrophenyl chromophore of Nph. This shift, from an absorbance maximum of 280 nm to one of 272 nm, has been documented in previous reports. The procedures utilized in this laboratory are described in full in a recent chapter [10]; briefly, a Hewlett-Packard diode-array spectrophotometer is utilized to collect multiple absorbance points in the

range of 284-324 nm. We take advantage of the averaging capacity of the computer to provide data smoothing as well as error estimates of the measurements. All experiments were conducted at 37°C.

Analysis of the structural relationships was carried out using the SYBYL program of Tripos Associates (St. Louis, MO, USA) and the coordinate sets 5PEP, 1LYA, and 2APR.

RESULTS AND DISCUSSION

Comparisons of kinetic data between different enzymes are dangerous, as it is possible that the mechanism of the systems might be different. In our studies, we have attempted to make all our evaluations within sets of closely related peptides. In addition, by our extensive variation of structure we have managed to identify at least some substrates that yield comparable optimal levels of activity for each enzyme, as will be detailed below. This indicates that the *intrinsic* mechanism and catalytic efficiency of each enzyme is similar, thus validating our analysis.

In the following discussion we will focus on the parameter k_{cat}/K_m, which is denoted the *specificity constant* for a given enzyme-substrate pair. This is a fundamental kinetic constant, and is preferable for the illustrations to be described. The use of the K_m parameter for comparing results between enzymes should be avoided, unless the true meaning of that constant has been defined for the enzymes in question.

The specificity constants for the three enzymes are listed in Table 1 for ten peptide substrates with changes in the P_3 and P_2 positions. The larger set of data, including all the peptides listed in Figure 3 and the kinetic parameters k_{cat} and K_m, will be reported separately (P. E. S., W. T. L., and C. R.-N., manuscripts in preparation).

Table 1. Specificity Constants for the Enzyme-Catalyzed Cleavage of Peptide Substrates of the Form Lys-Pro-*Xaa-Yaa*-Phe*Nph-Arg-Leu

	PORCINE PEPSIN	HUMAN CATHEPSIN D	*RHIZOPUS CHINENSIS* PROTEINASE
P_5-P_4-P_3-P_2-P_1*P_1'-P_2'-P_3'		k_{cat}/K_m $(M^{-1}s^{-1}) \times 10^{-6}$	
Lys-Pro-*Ala-Lys*-Phe*Nph-Arg-Leu	1.1	< 0.02	1.3
Lys-Pro-*Asp-Lys*-Phe*Nph-Arg-Leu	0.1	< 0.02	0.7
Lys-Pro-*Ser-Lys*-Phe*Nph-Arg-Leu	0.3	< 0.02	0.6
Lys-Pro-*Leu-Lys*-Phe*Nph-Arg-Leu	0.4	< 0.02	1.4
Lys-Pro-*Arg-Lys*-Phe*Nph-Arg-Leu	< 0.02	< 0.02	1.4
Lys-Pro-*Ala-Ala*-Phe*Nph-Arg-Leu	0.8	0.2	2.4
Lys-Pro-*Ala-Asp*-Phe*Nph-Arg-Leu	0.5	0.3	0.8
Lys-Pro-*Ala-Ser*-Phe*Nph-Arg-Leu	0.5	0.1	1.0
Lys-Pro-*Ala-Leu*-Phe*Nph-Arg-Leu	0.7	1.0	2.3
Lys-Pro-*Ala-Arg*-Phe*Nph-Arg-Leu	0.5	< 0.02	2.2

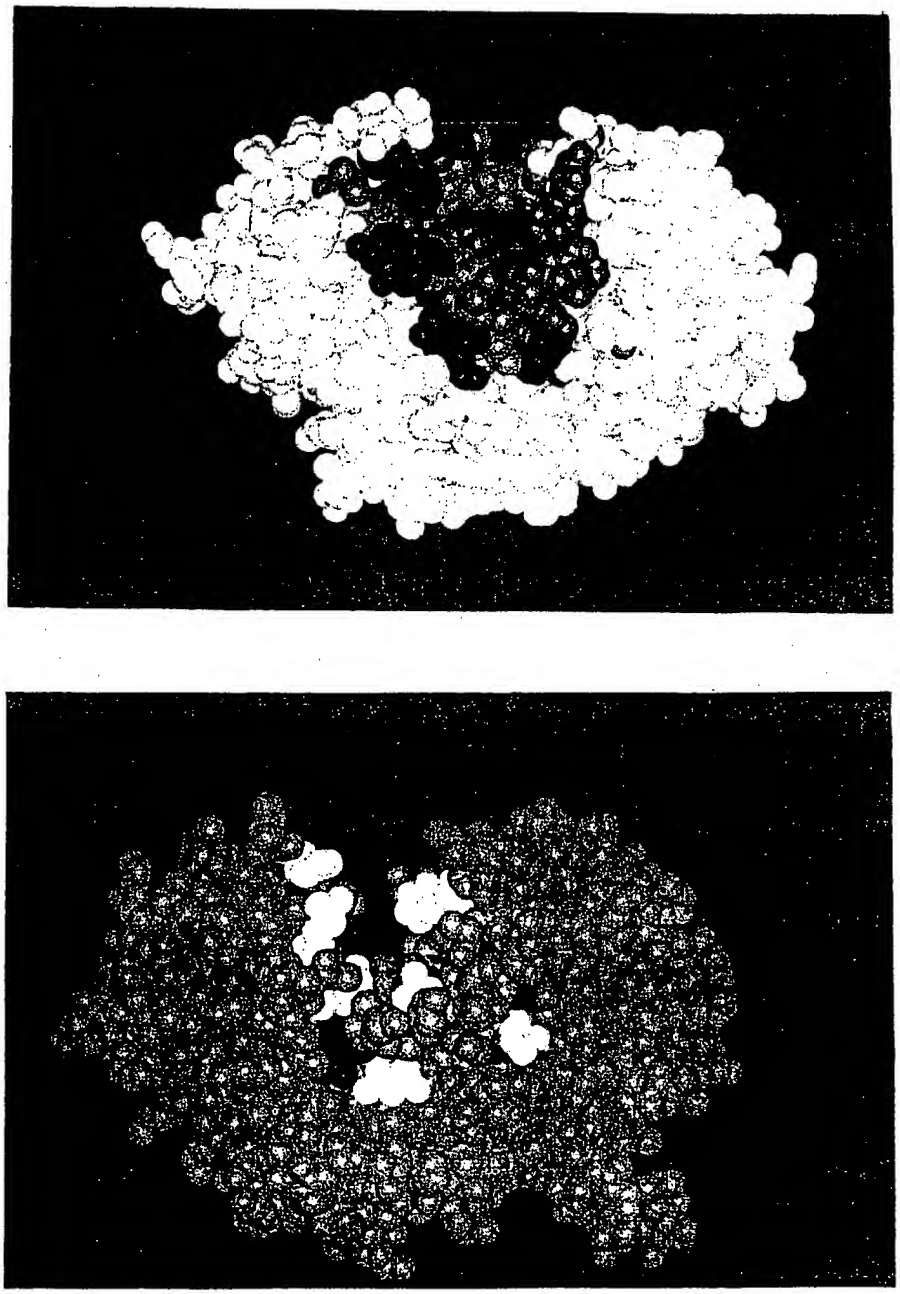


Figure 4. Representation of the acidic (Asp/Glu) and hydrophobic residue distribution within the active site clefts of (A) porcine pepsin and (B) human cathepsin D. Asp and Glu side chains are colored red in these space-filling diagrams. Hydrophobic residues are colored purple, and hydrophobic residues (e.g., Ser, Thr, Asn. Gln) are colored green. Residues outside the active site cleft are shown as cyan spheres. Both structures are oriented at the same angle. (See color plate preceding page 1.)

Compariso

In
seen in th
ences in th
the peptid
the peptid
at measura
similar rat
In
at nearly e
the peptid
with Ala-l
residue in
In
contrastin
substrate s
peptides is
chain subs
subsite wi
example, a
Leu, with
for pepsin.
On
unfavorabl
cleft, the c
Th
same; plac
hydrophob
In
one in wh
region of t
(Tyr_{75}, Ph
would not
could offer
pH of the a
When the
the P_3 posi
formation
Site-direct
For
position ar
position. E
region of t
287, than t
were later
cathepsin I
subsite pre
[14]. This
substrate a
Ov
of human c
or Glu resi

In examining these data, we will begin with several generalizations. First, it can be seen in the upper section (P_3 substitutions) that the three enzymes show significant differences in their specificity toward the five substrates with Lys in the P_2 position. Pepsin cleaves the peptides with Ala, Asp, Ser, and Leu in P_3 at roughly equivalent rates, but does not cleave the peptide containing Arg in the P_3 position; cathepsin D cleaves none of the five peptides at measurable rates; while *Rhizopus chinensis* proteinase cleaves all five peptides at roughly similar rates, and is the most efficient enzyme.

In the lower section it can be seen that all three enzymes cleave most of the peptides at nearly equivalent efficiencies; the *significant* exception is that cathepsin D does not cleave the peptide with Arg in the P_2 position. This result can be combined with the topmost peptide, with Ala-Lys in P_3-P_2, confirming that cathepsin D does not tolerate a positively-charged residue in the P_2 position.

In comparing all the information given in Table 1, we can see that there are two contrasting themes. On the one hand it is clear that some substitutions within the peptide substrate sequence are without (major) consequence; this helps to confirm that this *series* of peptides is interacting with the various enzymes in a common fashion. As long as the side chain substituted into the peptide can be accommodated within the corresponding enzyme subsite without penalty, the resulting catalytic efficiency is approximately equivalent. For example, all three enzymes cleave the peptide of sequence, Lys-Pro-Ala-Leu-Phe*Nph-Arg-Leu, with nearly equal k_{cat}/K_m values ($0.7 \times 10^6\ M^{-1}s^{-1}$, $1.0 \times 10^6\ M^{-1}s^{-1}$, and $2.3 \times 10^6\ M^{-1}s^{-1}$, for pepsin, cathepsin D, and *Rhizopus chinensis* proteinase, respectively).

On the other hand, when a substitution is made that leads to an energetically unfavorable interaction with the enzyme surface at that particular point in the active site cleft, the catalytic reaction is dramatically altered.

The source of the energetic penalty in both cases shown in Table 1 is essentially the same; placing a positively charged residue in a position where it is forced to interact with a hydrophobic subsite of the enzyme surface.

In the case of porcine pepsin, the only substrate that shows reduced cleavage is the one in which an Arg is present in the P_3 position. Examination of the structure of the S_3 region of the active site cleft of this enzyme reveals a preponderance of hydrophobic residues (Tyr_{75}, Phe_{111}, Ala_{115}). Thus, it is reasonable that the positively charged side chain of Arg would not interact favorably with this environment. The only residue in this vicinity that could offer an energetically beneficial interaction is the side chain of Glu_{13}; however, at the pH of the assay reported in Table 1 (pH 3.5), the side chain of this residue would be neutral. When the pH of the assay is increased, the rate of cleavage of the peptide containing Arg in the P_3 position increases, presumably due to loss of a proton from Glu_{13} and the consequent formation of an ion-pair interaction (C. R.-N. and B. M. D., unpublished observations). Site-directed mutagenesis at position 13 is now being employed to test this hypothesis.

For human cathepsin D, all substrates with a positively-charged residue in the P_2 position are cleaved at rates significantly below substrates with Ala, Asp, Ser, or Leu in that position. Examination of a molecular model of human cathepsin D [11] revealed that the S_2 region of that enzyme is more hydrophobic, due to the presence of a Met residue at position 287, than that of porcine pepsin, which has a Glu in the same position. These observations were later confirmed upon the publication of the three-dimensional structure of human cathepsin D by two groups [12,13]. Site-directed mutagenesis has been utilized to alter the subsite preferences of human cathepsin D by changing the Met at position 287 into a Glu [14]. This mutant enzyme cleaves the peptides containing Lys or Arg in the P_2 position of substrate at readily measurable rates [14].

Overall, the active site cleft of porcine pepsin is significantly more acidic than that of human cathepsin D, as illustrated in Figure 4, panels A and B. The large number of Asp or Glu residues within the active site cleft of porcine pepsin accounts for the rapid cleavage

of peptides such as those shown in Figure 3 and Table 1, where the net charge on each peptide varies from + 1-4 in the acidic pH range under consideration. We describe this phenomenum as a "general electrostatic effect", whereas the effect observed on titration of Glu_{13} in porcine pepsin would be described as a "specific electrostatic effect". We differentiate these effects based on the results generated with specific substitutions, such as those listed in Figure 3.

Finally, it can be seen in Table 1 that the proteinase from *Rhizopus chinensis* is the most versatile enzyme of the three compared, as it is able to cleave all the peptides listed at roughly comparable rates. In particular, it is of great interest that the *Rhizopus* enzyme is the only one that can cleave the peptide with Arg in the P_3 position at pH 3.5. Comparing the active site structures of the three enzymes, it is clear that *Rhizopus chinensis* aspartic proteinase is unique in the presence of two *additional* acidic residues: Asp_{30} and Asp_{77} (using the pepsin numbering system). Asp_{30} is located at the interface of the S_1 and S_3 subsites, on the N-terminal domain, while Asp_{77} is located on the mobile "flap" that hangs over the active site cleft. Asp_{77} can interact with residues in several positions, including at least P_1 and P_2. In addition, *Rhizopus chinensis* aspartic proteinase has Glu at position 13, while the residue corresponding to 287 is a Gly. Thus, the *Rhizopus* enzyme is identical to pepsin at position 13, but is more like human cathepsin D at position 287. The lack of an acidic residue at 287 is compensated for by the additional acidic side chains at 30 and 77. This leads to the *Rhizopus* enzyme demonstrating the most flexibility in accommodation of substrates, cleaving all the 10 peptides offered to it in this study.

Again, the unique residues in the *Rhizopus chinensis* aspartic proteinase are the targets of site-directed mutagenesis (T. L. and B. M. D., unpublished observations) and will be the subject of future publications from this laboratory.

In summary, the comparison presented here illustrates the dominant role of a few residues in determining the "secondary specificity" of the aspartic proteinases [5]. While the overall conformation of the three enzymes are extremely similar, the specific amino acid side chains lining the pockets of the various subsites of the enzymes are variable. The residues can provide either an attractive or a repulsive force when interacting with a complementary side chain of a ligand.

The differences in specificity revealed by the comparison of cleavage rates of the sets of peptides presented in Figure 3 and Table 1 provide information that could be of value in the design of selective inhibitors for each enzyme. Conversely, substituents could be chosen, based on this information, to design a compound that would *avoid* interaction with a specific enzyme. It will be of interest to determine the patterns of specificity of other members of this class. Finally, the data obtained by analysis of the set of peptides given in Figure 3 provides an important database of information that is essential to the analysis of the effects of site-directed mutagenesis of various members of the aspartic proteinase class.

REFERENCES

1. *Acid Proteases: Structure, Function, and Biology* (J. Tang, Ed.), Advances in Experimental Medicine and Biology, Volume 95, Plenum Press, New York, (1977).
2. *Aspartic Proteinases and Their Inhibitors* (V. Kostka, Ed.), Walter de Gruyter, New York, (1985); *Structure and Function of the Aspartic Proteinases* (B. M. Dunn, Ed.), Advances in Experimental Medicine and Biology, Volume 306, Plenum Press, New York, (1991).
3. J. C. Powers, A. D. Harley, and D. V. Myers, *in: Acid Proteases: Structure, Function, and Biology* (J. Tang, Ed.), Advances in Experimental Medicine and Biology, Volume 95, Plenum Press, New York, pp. 141-157 (1977).
4. V. K. Antonov, *in: Acid Proteases: Structure, Function, and Biology* (J. Tang, Ed.), Advances in Experimental Medicine and Biology, Volume 95, Plenum Press, New York, pp. 179-198 (1977).
5. B.M. Dunn, B. Kammermann, and K.R. McCurry, Anal. Biochem., 138: 68-73 (1984).

6. J. S. Fruton, Adv. Enzymol. 44:1-36 (1976).
7. T. Hofmann, B. Allen, M. Bendiner, M. Blum, and A. Cunningham, Biochem., 27:1140-1146 (1988).
8. M. Balbaa, A. Cunningham, and T. Hofmann, Archives B. B., 306:297-303 (1993).
9. J. Pohl and B.M. Dunn, Biochem., 27: 4827-4834 (1988).
10. B.M. Dunn, P.E. Scarborough, and W. Swietnicki, *in: Peptide Synthesis and Purification Protocol*, (B.M. Dunn and M.W. Pennington, Eds.), Humana Press, Clifton, N.J., in press (1993).
11. P.E. Scarborough, K. Guruprasad, T.L. Blundell, G.R. Richo, and G.E. Conner, Protein Science, 2: 264-276 (1993).
12. P. Metcalf and M. Fusek, EMBO J. 12:1293-1302.
13. E. T. Baldwin, T. N. Bhat, S. Gulnik, M. Hosur, R. C. Sowder II, R. E. Cachau, J. Collins, A. M. Silva, and J. W. Erickson, Proc. Natl. Acad. Sci. (USA) 90:6796-6800.
14. P.E. Scarborough and B.M. Dunn, Protein Engineering, in press (1994).

ELSEVIER

Making recombinant proteins in animals – different systems, different applications

Michael K. Dyck, Dan Lacroix, François Pothier and Marc-André Sirard

Centre de Recherche en Biologie de la Reproduction Dépt des Sciences Animals, Pavillon Paul Comtois, Cité Universitaire, Université Laval, Sainte-Foy, Québec, Canada, G1K 7P4

Transgenic animal bioreactors represent a powerful tool to address the growing need for therapeutic recombinant proteins. The ability of transgenic animals to produce complex, biologically active recombinant proteins in an efficient and economic manner has stimulated a great deal of interest in this area. As a result, genetically modified animals of several species, expressing foreign proteins in various tissues, are currently being developed. However, the generation of transgenic animals is a cumbersome process and remains problematic in the application of this technology. The advantages and disadvantages of different transgenic systems in relation to other bioreactor systems are discussed.

The biotechnology industry is currently experiencing an extreme shortage of manufacturing capacity for recombinant therapeutic proteins [1]. As a result, a growing number of biological systems are being evaluated for the production of these valuable proteins. Although some have been used for many years, others are relatively new and still experimental. Factors such as scale-up, total annual production, speed of production set-up, post-translational modifications and regulatory issues come into play in choosing the system that is most suitable for any given protein target [2–6] (Fig. 1). This review discusses different systems currently being considered and applied for recombinant protein production and focuses on the use of transgenic animals for this purpose.

Recombinant protein production platforms

Bacteria have proven useful as bioreactors because they are grown easily at any scale. However, they are limited in their ability to perform the post-translational protein modifications necessary for many targets [7–9]. Certain eukaryotic systems, such as yeast, filamentous fungi and unicellular algae, can be scaled-up with relative ease [10–13] and are capable of post-translational modifications. However, these systems are often limited by their ability to duplicate human patterns of protein processing and can thus yield recombinant products with undesirable properties, such as immunogenicity or lack of activity. Insect cell systems are commonly used at the laboratory scale and some systems offer adequate production yields [14], but they have unique glycosylation patterns and the baculovirus system is more appropriate for laboratory scale production. Metazoa cell culture systems have also been used as bioreactors but are expensive to maintain and difficult to scale-up. Mammalian cells in particular can perform complex post-translational modifications, although the costs associated with scaling these systems up for mass-production purposes are extremely high [15]. Transgenic plants [16,17], animals [18–20] and insects [21] have a potentially large production capacity at lower costs than mammalian cell culture but involve relatively slow production set-up and have yet to cross many regulatory hurdles.

Direct comparison of the production costs associated with these different systems can be difficult because of the lack of data on protein yield, purification rates and production scale, particularly for new systems. The specific recombinant protein being produced also has a major role in defining each of these factors and numbers can vary according to specific costs. Capital and production costs favour transgenic animals over mammalian cell culture. Building a large-scale (10 000 l bioreactor) manufacturing facility for mammalian cells takes 3–5 years and costs US$250–500 million, whereas a transgenic farm with a single purification facility should not cost more than US$80 million, probably less. As seen in Table 1, production costs are substantially lower for transgenics than for cell culture (presented by H.L. Levine, 2002 BIO International Biotechnology Convention and Exhibition, 9–12 June, Toronto, Ontario, Canada). However, once the raw material has been produced, validated purification processes costs are similar regardless of the production systems used and bring total estimated costs-of-goods

Table 1. Comparative estimated production COGS between cell culture and transgenics[a]

Production scale	System	Production, COGS (dollars gram^{-1})
50 kg year^{-1}	Cell culture	147
	Transgenics	20
300 kg year^{-1}	Cell culture	48
	Transgenics	6

Abbreviation: COGS, costs-of-goods.
[a]Presented by H.L. Levine, 2002 BIO International Biotechnology Convention and Exhibition, 9–12 June, Toronto, Ontario, Canada.

Corresponding author: Marc-André Sirard (Marc-Andre.Sirard@crbr.ulaval.ca).

http://tibtec.trends.com 0167-7799/$ - see front matter doi:10.1016/S0167-7799(03)00190-2

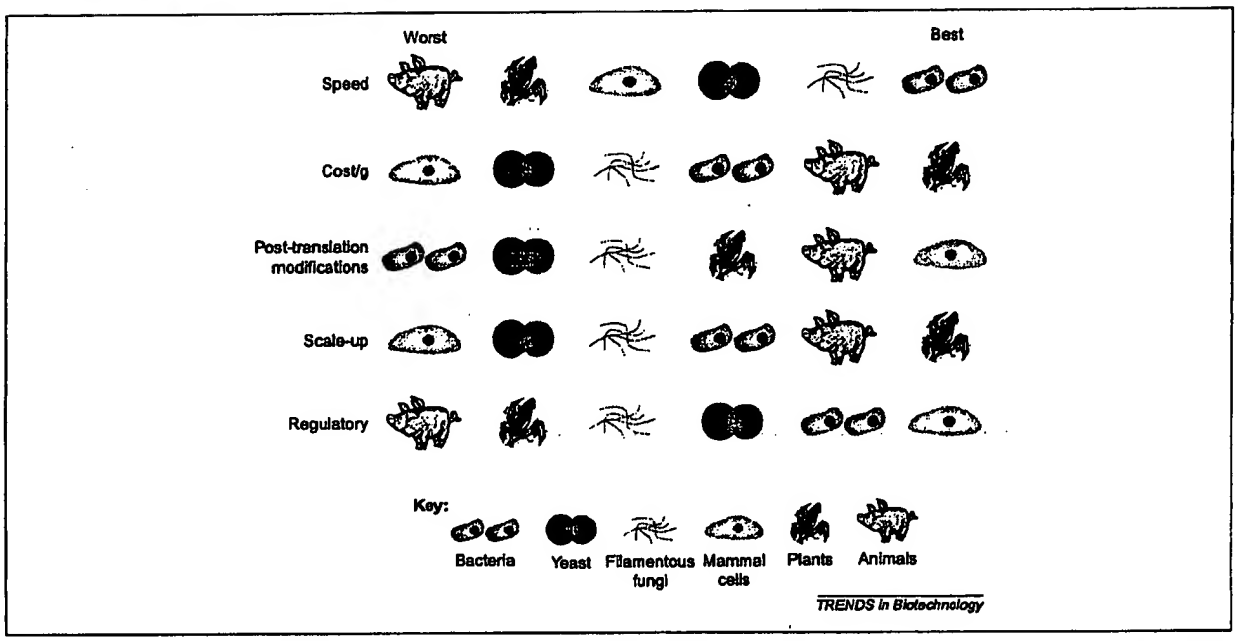


Fig. 1. Relative merits of different systems. Speed, 'gene to production' time; Cost/g, total cost of goods; scale-up, ease and speed; regulatory, accumulated products approval history.

(COGS) closer between cell culture and transgenics. In fact, an estimation of total COGS which include capital, production and purification costs gives values of US$942 gram^{-1} for cell culture and US$679–703 gram^{-1} for transgenics for a production scale of 50 kg year^{-1} (presented by H.L. Levine, BIO 2002). Thus, although the gap is not as large when purification is factored in, transgenics still show a financial advantage over cell culture even when all costs are taken into account.

This type of evaluation demonstrates the economic efficiency of producing recombinant proteins using transgenic systems compared with cell culture bioreactors. This, coupled with the fact that transgenic animals are well equipped to perform all of the complex post-translational modifications necessary to render some proteins biologically active, has driven interest in developing transgenic livestock as bioreactors to produce valuable recombinant proteins in their bodily fluids.

Transgenic animal bioreactors

The mammary gland has generally been considered the tissue of choice to express valuable recombinant proteins in transgenic animal bioreactors because milk is easily collected in large volumes. As a result, a great deal of effort has been made to produce transgenic bioreactors with the traditional 'dairy' species, such as sheep, goats and cows [22]. Foreign proteins are commonly reported to be expressed into transgenic milk at rates of several grams per litre [23]. However, the production of proteins in milk is limited by the relatively long interval from birth to first lactation encountered with domestic livestock, the discontinuous nature of the lactation cycle and the substantial time and material investments required to produce transgenic dairy animals [24]. Transgenic rabbits and pigs expressing foreign proteins in their mammary glands have been produced to address this problem but milk production rates and the number of animals needed to produce adequate amounts of protein can be limiting [23]. In addition, certain bioactive proteins produced in milk can have adverse affects on the animal's health. This is particularly true when they are produced at high concentrations and the protein can be reabsorbed. This limits the use of this type of recombinant protein production system to inactive or non-interfering proteins [25].

The use of transgenic eggs for large-scale production of recombinant proteins is another method being contemplated. Interest in this system is driven by the fact that a single hen can produce an impressive number of eggs (up to 330 eggs year^{-1}) and egg white naturally contains ~4 g of protein [26]. Despite its potential however, this system has been hampered by the lack of an efficient system of transgenesis in poultry.

Other forms of collectable bodily fluids that could be used for the production of foreign proteins in transgenic animals are being considered. The possibility of isolating foreign proteins from the blood of transgenic pigs has been explored and pigs producing human hemoglobin in their own circulatory system have been produced [27]. In principle, the human component of the pigs' blood was to be used as a blood substitute, but the similarities between the porcine and human blood components made isolation of the human hemoglobin arduous. Blood is a less than ideal fluid for protein production because harvesting is invasive and bioactive proteins could affect the animal's health to the point of making it impractical. The idea of using the bladder as a bioreactor by engineering urethelium production and secretion of a foreign protein into the urine has also been explored [28–31]. The limiting factor with bladder production of proteins has been yield.

Although the bladder epithelium does secrete proteins, the rates are minimal, and thus protein production rates with this system are extremely low.

The seminal fluid of the male ejaculate has also been considered as a site for recombinant protein secretion in transgenic animals [32]. Of particular interest is the pig because the boar's male accessory sex glands possess many characteristics that make them appropriate for the production of recombinant proteins, including: a large capacity for protein production; protein production is continuous throughout the reproductive life of the animal; the ability to perform the complex post-translational modifications. Pig semen contains 30 mg of protein per ml [33] and boars can produce 200–300 ml of semen [34] for a total of 6–9 g of protein per ejaculate. The collection and handling of boar semen is a well-established process, performed on a large scale at swine artificial insemination units worldwide. Also of interest is the fact that protein secretion by these tissues is uniquely exocrine, minimising the risk of a biologically active recombinant protein upsetting the host's own physiology.

The generation of transgenic pig bioreactors producing foreign proteins in their semen will initially be limited by our lack of knowledge regarding the regulatory sequences and promoters to drive expression of proteins into the male sex glands. Therefore the isolation and characterization of these sequences is necessary. Given that the family of proteins referred to as spermadhesins are the major protein component of boar seminal fluid [35], expression of recombinant protein coding sequences under the control of the promoter regions of these genes in transgenic boars will provide an indication of the production capacity of this bioreactor system.

The raw potential for producing valuable proteins with transgenic animals seems apparent. However, the purification of these proteins from their source, whether milk, eggs or semen, is still a hurdle to be overcome and creates, often undefined, regulatory issues. Isolation of recombinant proteins from milk is complicated by the presence of micelles and fat globules [36]. Purification challenges inherent to the complex composition of the egg could also be problematic. However, for semen, once the sperm has been removed from the seminal fluid, protein purification can be performed using methods previously established for milk. Another aspect to consider when producing proteins in the tissues of transgenic animals is the ability of the tissues to execute complex post-translational modifications. This process is different from protein to protein and might also vary from tissue to tissue.

Generating transgenic animals

Although transgenic animal bioreactors represent a powerful means of producing recombinant proteins, the generation of transgenic domestic animals is difficult and often considered a barrier to their application. The technique that has been the most successful in producing transgenic animals is the microinjection of DNA into the pronuclei of fertilised oocytes [37,38]. The efficiency of transgenesis in large domestic animals varies but is generally considered to approach 1.0% [39]. This degree of inefficiency, coupled with the extended gestation and high maintenance costs of farm animals, makes the production of certain species of transgenic livestock by this means time-consuming and expensive. The nature of avian reproductive systems makes this form of gene transfer impossible in poultry. Furthermore, the unpredictability of the site and rate of transgene integration in the host genome and the resulting variation in transgene expression because of position effects have also proved problematic [40]. These limitations have driven the search for alternate modes of transgenesis, resulting in several unique approaches to gene transfer.

Retroviruses represent a natural system capable of efficiently introducing foreign DNA into animal cells [41]. For gene transfer purposes, viral gene sequences are deleted from the organism and replaced with a transgene. The redesigned retroviruses are introduced into developing embryos to facilitate the transfer of the foreign DNA into an animal. The use of retroviral vectors for transgenesis is unique in that only a single copy of the transgene is integrated in the host genome and the virus can be introduced into oocytes or embryos at various stages. A less than ideal aspect of retroviral vectors is that they are limited in the size of constructs that they can carry [42,43]. Furthermore, founder animals are generally mosaic and the genes are not always expressed in the second generation. Despite this, retroviruses have been used to successfully produce transgenic mice [44,45] and viral integration of recombinant sequences into bovine embryos to produce transgenic calves has been reported [46,47].

The use of motile sperm as vectors to introduce foreign DNA into oocytes has stimulated great interest. The first reported use of sperm as DNA vectors involved the incubation of washed mouse spermatozoa in the presence of DNA fragments, leading to the production of transgenic mice when these spermatozoa were used for *in vitro* fertilisation [48]. The same group has reported of the production of transgenic pigs with this technique [49–51]. Attempts to increase DNA binding to the sperm with DNA–liposome complexes [52], electroporation [53,54] or with the aid of antibodies [55] have also been explored. This form of transgenesis is enticing because it requires neither specialized equipment nor a high level of expertise. However, the technique is continually confounded by the limited ability of the host's genome to integrate foreign DNA presented in this manner. As a result, the sperm-mediated production of transgenic animals has been difficult to reproduce [56] and generally results in mosaic animals [55].

A twist on the use of sperm as DNA carriers involves manipulating the cells responsible for spermatogenesis referred to as spermatogonia, rather than the sperm themselves [57]. A series of publications by Brinster and colleagues brought attention to the potential of being able to recover spermatogonial stem cells, genetically manipulate them *in vitro*, and transplant the cells into a recipient testis [58–60]. The recipient animals act as vectors, producing male gametes originating from the genetically modified spermatogonia. Resulting transgenic offspring would harbour the gene introduced into the male stem cells *in vitro*. Processes for transplanting testis cells from one male to another, as well as culturing spermatogonial cells have been established in the mouse [61].

However, development of this technology in livestock has been limited to the manipulation of the pig male stem cells *in vivo* [62,63].

Embryonic stem (ES) cells and primordial germ (PG) cells provide another medium for the production of transgenic animals and represent the primary means of gene transfer for poultry. The ability to isolate, maintain and manipulate pluripotent ES or PG cells is a powerful research tool [64,65]. Genetically modified cells are injected into developing embryos to produce chimeric animals. If the modified cell line contributes to the gonads of the chimeric animal and participates in the production of sperm and oocytes, resulting offspring will include a certain percentage of transgenic progeny. The advantages of ES cells as a mode of gene transfer include: (1) ES cells can be transformed *in vitro* with foreign DNA and screened before being used to produce chimerics; (2) the site of transgene integration in the genome can be controlled by homologous recombination to replace existing genes. Reviews of the literature indicate that the production of chimeric animals with ES or PG cell technology has been applied successfully in mice [66,67] rabbits [68], pigs [69], cattle [70] and poultry [71]. However, transmission of the ES or PG genome into the gametes to produce transgenic offspring from a chimeric animal has only been efficiently achieved in mice [72] and chickens [73].

The production of transgenic animals by nuclear transfer offers the same primary advantage as ES cells, in that genetic manipulations can be performed on cell lines *in vitro*. The nuclear material from these modified cells is then transferred into the cytoplasm of a recipient cell from which the genetic material has been removed. The resulting entity is exposed to an activation process, which if successful, causes it to divide and develop into an animal. Therefore, characterized cell lines in which the desired expression patterns of a particular transgene have been established can be used as nuclear donors. The resulting animals are genetic copies of the cells manipulated in culture and therefore carry the transgene of interest.

The ability to produce 'clones' by nuclear transfer holds great potential for the area of genetic manipulation of livestock and has been used to produce transgenic sheep [74], cattle [75,76] and pigs [77]. Recently, cloned calves that harbour an artificial mammalian chromosome have been produced using these same procedures [78]. However, this technology is still plagued by low clone viability, with most dying during gestation or soon after birth. Surviving offspring also exhibit increased birth weights and pathological features [79,80]. Therefore, identification and elimination of the factors resulting in these adverse affects is necessary before this technique can be universally applied for this purpose.

Pronuclear microinjection, despite its limitations, remains the most straightforward and consistently successful means of gene transfer for most species. Preparation of a transgene for microinjection requires little beyond the techniques necessary to produce any form of DNA construct. The manipulations involved in microinjection, although challenging, are no more cumbersome than those required for other techniques. Poor embryo survival to term, low transgene integration and the unpredictability of transgene behaviour is problematic and has lead to the search for alternative gene transfer strategies. However, none of the alternatives to date has done so without burdening the transgenic animal production system with additional pitfalls. Furthermore, for reproductively efficient species, including mice, rabbits and pigs, this inefficiency is less prohibitive than for less prolific species, such as goats, sheep and cattle.

Commercial application of transgenic animal bioreactors

The use of transgenic animals for protein production in a research environment is generally performed without constraints but can become limited when considering commercial applications. Many aspects of the more recent approaches described above have been patented for agricultural or biomedical applications. For example, to express a particular protein in the mammary gland, a functional promoter for this tissue is needed. Unfortunately, all the regulatory sequences for this purpose are currently covered by patent limitations. In addition, if the product one wishes to produce in milk has a known DNA sequence, chances are the protein and its use as a therapeutic are also patented. Furthermore, if nuclear transfer or cloning is a part of the process, a myriad of patents have been granted for various aspects of this procedure. Currently, several patent holders maintain that they possess a valid patent to clone and freedom to operate, so it might take a few years and a certain degree of litigation to resolve the individual validity of these claims. Also, because many patents are based on slightly modified technical approaches, it might be difficult to determine which techniques are used to obtain a final product. Regulatory processes associated with the commercialization of a given product will require the description of detailed procedures and help in the enforcement of intellectual property rights.

Other concerns surrounding these technologies include the ethical and environmental aspects of transgenesis. Integration of a transgene into the genome might disturb endogenous gene expression either in the first generation or when homozygote transgenic animals are required. In the past, certain gene transfer studies have resulted in affected or sick animals [25,81] and these experiments were terminated. Any genetic manipulation resulting in animal suffering would not be acceptable to scientists, the public or regulatory agencies.

The environmental issues associated with genetically modified organisms that have caused the most public outcry, including food and ecosystem contamination or threats to biodiversity, are generally more problematic in transgenic plants than animals. Most of the applications of transgenic animals are related to biomedical applications and therefore does not include entry of the animals into the food chain. Unlike plants, there is little chance of transgenic domestic animals entering the wild and mating with feral species. In the case of biodiversity, people working with traditional genetic selection procedures in farm animals have already taken measures to preserve some rare breeds. However, theoretically, adding a transgene to a population actually increases biodiversity

unless, of course, the transgenic lines replace other lines that are less resistant to a disease.

Finally, social issues and concerns surrounding transgenics have been raised and continue to propagate, putting the science of biotechnology under unprecedented scrutiny. Many of the concerns over the ethics of using transgenic animals, such as the safety of these products or the health of the genetically modified animal, can be addressed in a logical and scientific manner. However, there remain other less tangible social perceptions that might be very difficult to address. For example, the fact that a recombinant protein has been isolated from the urine of a transgenic animal might carry a very negative public perception and impede the marketability of that product. Those of us working with transgenic animals cannot ignore these concerns and will be forced to demonstrate the advantages and safety of products derived from these animal populations, but this form of social repugnancy could prove insurmountable for certain products.

Conclusion

The production of therapeutic proteins in transgenic animals continues to advance, with products in clinical trials or late-stage pre-clinical development, and so the future of this technology looks promising. However, current methods of generating transgenic animal founders are relatively inefficient and time-consuming, and attempts to improve transgenesis by various methods have had limited success. The inefficiency of transgenesis in the dairy species, as well as certain innate disadvantages of lactation, has prompted interest in expressing foreign proteins in various tissues of more prolific species. Although less well-established, results to date indicate that the eggs, urine and semen of transgenic animals could prove to be viable alternatives to mammary gland-based systems.

References

1 Dove, A. (2002) Uncorking the biomanufacturing bottleneck. *Nat. Biotechnol.* 20, 777–779
2 Werner, R.G. *et al.* (1998) Appropriate mammalian expression systems for biopharmaceuticals. *Arzneim.-Forsch. Drug Res.* 48, 870–880
3 Fischer, R. *et al.* (1999) Towards molecular farming in the future: moving from diagnostic protein and antibody production in microbes to plants. *Biotechnol. Appl. Biochem.* 30, 101–108
4 Verma, R. *et al.* (1998) Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems. *J. Immunol. Methods* 216, 165–181
5 Morton, C.L. and Potter, P.M. (2000) Comparison of *Escherichia coli*, Saccharomyces cerevisiae, *Pichia pastoris*, *Spodoptera frugiperda*, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase. *Mol. Biotechnol.* 16, 193–202
6 Bulleid, N.J. *et al.* (2000) Recombinant expression systems for the production of collagen. *Biochem. Soc. Trans.* 28, 350–353
7 Balbas, P. (2001) Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. *Mol. Biotechnol.* 19, 251–267
8 Collet, J.F. and Bardwell, J.C. (2002) Oxidative protein folding in bacteria. *Mol. Microbiol.* 44, 1–8
9 Swartz, J.R. (2001) Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol.* 12, 195–201
10 Conesa, A. *et al.* (2001) The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genet. Biol.* 33, 155–171
11 Cregg, J.M. *et al.* (2000) Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* 16, 23–52
12 Punt, P.J. *et al.* (2002) Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20, 200–206
13 Fuhrmann, M. *et al.* (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. *Plant J.* 19, 353–361
14 Farrell, P.J. *et al.* (1999) Transformed Lepidopteran insect cells: new sources of recombinant human tissue plasminogen activator. *Biotechnol. Bioeng.* 64, 426–433
15 Andersen, D.C. and Krummen, L. (2002) Recombinant protein expression for therapeutic applications. *Curr. Opin. Biotechnol.* 13, 117–123
16 Daniell, H. (2002) Molecular farming for health – production of antibodies, biopharmaceuticals and edible vaccines in plants. In *Healthcare & Pharmaceutical Briefing - Life Sciences Technology* (Briefings, W.M.S.B., ed.), pp. 1–8, World Markets Research Center, London, UK
17 Giddings, G. *et al.* (2000) Transgenic plants as factories for biopharmaceuticals. *Nat. Biotechnol.* 18, 1151–1155
18 Das, R.C. (2001) Production of therapeutic proteins from transgenic animals. In *American Biotechnology Laboratory* (Vol. Feb), pp. 60–64, International Science Communications, Shelton, Connecticut, USA
19 Harvey, A.J. *et al.* (2002) Expression of exogenous protein in the egg white of transgenic chickens. *Nat. Biotechnol.* 20, 396–399
20 Rudolph, N.S. (1999) Biopharmaceutical production in transgenic livestock. *Trends Biotechnol.* 17, 367–374
21 Tomita, M. *et al.* (2003) Transgenic silkworms produce recombinant human type III procollagen in cocoons. *Nat. Biotechnol.* 21, 52–56
22 Eyestone, W.H. (1999) Production and breeding of transgenic cattle using *in vitro* embryo production technology. *Theriogenology* 51, 509–517
23 Houdebine, L.M. (1995) The production of pharmaceutical proteins from the milk of transgenic animals. *Reprod. Nutr. Dev.* 35, 609–617
24 Wall, R.J. *et al.* (1997) Transgenic dairy cattle: genetic engineering on a large scale. *J. Dairy Sci.* 80, 2213–2224
25 Massoud, M. *et al.* (1996) The deleterious effects of human erythropoietin gene driven by the rabbit whey acidic protein gene promoter in transgenic rabbits. *Reprod. Nutr. Dev.* 36, 555–563
26 Harvey, A.J. *et al.* (2002) Expression of exogenous protein in the egg white of transgenic chickens. *Nat. Biotechnol.* 20, 396–399
27 Swanson, M.E. *et al.* (1992) Production of functional human hemoglobin in transgenic swine. *Biotechnology* 10, 557–559
28 Kerr, D.E. *et al.* (1998) The bladder as a bioreactor: urothelium production and secretion of growth hormone into urine. *Nat. Biotechnol.* 16, 75–79
29 Zbikowska, H.M. *et al.* (2002) The use of the uromodulin promoter to target production of recombinant proteins into urine of transgenic animals. *Transgenic Res.* 11, 425–435
30 Zbikowska, H.M. *et al.* (2002) Uromodulin promoter directs high-level expression of biologically active human alpha1-antitrypsin into mouse urine. *Biochem. J.* 365, 7–11
31 Ryoo, Z.Y. *et al.* (2001) Expression of recombinant human granulocyte macrophage-colony stimulating factor (hGM-CSF) in mouse urine. *Transgenic Res.* 10, 193–200
32 Dyck, M.K. *et al.* (1999) Seminal vesicle production and secretion of growth hormone into seminal fluid. *Nat. Biotechnol.* 17, 1087–1090
33 Setchell, B.P. and Brooks, D.E. (1988) Anatomy, vasculature, innervation, and fluids of the male reproductive tract. In *The Physiology of Reproduction* (Vol. 1) (Knobil, E.A.N.J., ed.), pp. 753–836, Raven Press
34 Anderson, L.L. (1993) Pigs. In *Reproduction in Farm Animals* (Hafez, E.S.E., ed.), pp. 324–345, Lea & Febriger
35 Solis, D. *et al.* (1998) Binding of mannose-6-phosphate and heparin by boar seminal plasma PSP-II, a member of the spermadhesin protein family. *FEBS Lett.* 431, 273–276
36 Wilkins, T.D. and Velander, W. (1992) Isolation of recombinant proteins from milk. *J. Cell. Biochem.* 49, 333–338
37 Gordon, J.W. *et al.* (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. U. S. A.* 77, 7380–7384
38 Hammer, R.E. *et al.* (1985) Production of transgenic rabbits, sheep and pigs by microinjection. *Nature* 315, 680–683
39 Pinkert, C.A. and Murray, J.D. (1999) Transgenic farm animals. In *Transgenic Animals in Agriculture* (Murray, J.D. *et al.*, eds), pp. 1–18, CABI

40 Clark, A.J. *et al.* (1994) Chromosomal position effects and the modulation of transgene expression. *Reprod. Fertil. Dev.* 6, 589–598
41 Varmus, H. (1988) Retroviruses. *Science* 240, 1427–1435
42 Friedrich, G. and Soriano, P. (1993) Insertional mutagenesis by retroviruses and promoter traps in embryonic stem cells. *Methods Enzymol.* 225, 681–701
43 Cepko, C.L. *et al.* (2000) Lineage analysis with retroviral vectors. *Methods Enzymol.* 327, 118–145
44 van der Putten, H. *et al.* (1985) Efficient insertion of genes into the mouse germ line via retroviral vectors. *Proc. Natl. Acad. Sci. U. S. A.* 82, 6148–6152
45 Rubenstein, J.L. *et al.* (1986) Introduction of genes into preimplantation mouse embryos by use of a defective recombinant retrovirus. *Proc. Natl. Acad. Sci. U. S. A.* 83, 366–368
46 Haskell, R.E. and Bowen, R.A. (1995) Efficient production of transgenic cattle by retroviral infection of early embryos. *Mol. Reprod. Dev.* 40, 386–390
47 Chan, A.W. *et al.* (1998) Transgenic cattle produced by reverse-transcribed gene transfer in oocytes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14028–14033
48 Lavitrano, M. *et al.* (1989) Sperm cells as vectors for introducing foreign DNA into eggs: genetic transformation of mice. *Cell* 57, 717–723
49 Lavitrano, M. *et al.* (1997) Sperm-mediated gene transfer: production of pigs transgenic for a human regulator of complement activation. *Transplant. Proc.* 29, 3508–3509
50 Lavitrano, M. *et al.* (2002) Efficient production by sperm-mediated gene transfer of human decay accelerating factor (hDAF) transgenic pigs for xenotransplantation. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14230–14235
51 Lavitrano, M. *et al.* (2003) Sperm mediated gene transfer in pig: selection of donor boars and optimization of DNA uptake. *Mol. Reprod. Dev.* 64, 284–291
52 Bachiller, D. *et al.* (1991) Liposome-mediated DNA uptake by sperm cells. *Mol. Reprod. Dev.* 30, 194–200
53 Gagne, M.B. *et al.* (1991) Electroporation of bovine spermatozoa to carry foreign DNA in oocytes. *Mol. Reprod. Dev.* 29, 6–15
54 Horan, R. *et al.* (1991) Association of foreign DNA with porcine spermatozoa. *Arch. Androl.* 26, 83–92
55 Chang, K. *et al.* (2002) Effective generation of transgenic pigs and mice by linker based sperm-mediated gene transfer. *BMC Biotechnol.* 2, 5
56 Brinster, R.L. *et al.* (1989) No simple solution for making transgenic mice. *Cell* 59, 239–241
57 Brinster, R.L. (2002) Germline stem cell transplantation and transgenesis. *Science* 296, 2174–2176
58 Brinster, R.L. and Zimmermann, J.W. (1994) Spermatogenesis following male germ-cell transplantation. *Proc. Natl. Acad. Sci. U. S. A.* 91, 11298–11302
59 Brinster, R.L. and Avarbock, M.R. (1994) Germline transmission of donor haplotype following spermatogonial transplantation. *Proc. Natl. Acad. Sci. U. S. A.* 91, 11303–11307
60 Dym, M. (1994) Spermatogonial stem cells of the testis. *Proc. Natl. Acad. Sci. U. S. A.* 91, 11287–11289
61 Brinster, R.L. and Nagano, M. (1998) Spermatogonial stem cell transplantation, cryopreservation and culture. *Semin. Cell Dev. Biol.* 9, 401–409
62 Kim, J.H. *et al.* (1997) Development of a positive method for male stem cell-mediated gene transfer in mouse and pig. *Mol. Reprod. Dev.* 46, 515–526
63 Honaramooz, A. *et al.* (2003) Use of adeno-associated virus for trasfection of male-germ cells for transplantation in pigs. *Theriogenology* 59, 536
64 Fassler, R. *et al.* (1995) Knockout mice: how to make them and why. The immunological approach. *Int. Arch. Allergy Immunol.* 106, 323–334
65 Galli-Taliadoros, L.A. *et al.* (1995) Gene knock-out technology: a methodological overview for the interested novice. *J. Immunol. Methods* 181, 1–15
66 Evans, M.J. and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156
67 Martin, G.R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 78, 7634–7638
68 Giles, J.R. *et al.* (1993) Pluripotency of cultured rabbit inner cell mass cells detected by isozyme analysis and eye pigmentation of fetuses following injection into blastocysts or morulae. *Mol. Reprod. Dev.* 36, 130–138
69 Wheeler, M.B. (1994) Development and validation of swine embryonic stem cells: a review. *Reprod. Fertil. Dev.* 6, 563–568
70 Iwasaki, S. *et al.* (2000) Production of live calves derived from embryonic stem-like cells aggregated with tetraploid embryos. *Biol. Reprod.* 62, 470–475
71 Petitte, J.N. *et al.* (1990) Production of somatic and germline chimeras in the chicken by transfer of early blastodermal cells. *Development* 108, 185–189
72 Anderson, G.B. (1999) Embryonic stem cells in agriculture. In *Transgenic Animals in Agriculture* (Murray, J.D. *et al.*, eds), pp. 59–66, CABI
73 Bacon, L.D. *et al.* (2002) Identrification and evaluation of major histocompatibility complex antigens in chicken chimeras and their relationship to germline transmission. *Poultry Sci.* 81, 1427–1438
74 Schnieke, A.E. *et al.* (1997) Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science* 278, 2130–2133
75 Cibelli, J.B. *et al.* (1998) Transgenic bovine chimeric offspring produced from somatic cell-derived stem-like cells. *Nat. Biotechnol.* 16, 642–646
76 Cibelli, J.B. *et al.* (1998) Cloned transgenic calves produced from nonquiescent fetal fibroblasts. *Science* 280, 1256–1258
77 Lai, L. *et al.* (2002) Transgenic pig expressing the enhanced green fluorescent protein produced by nuclear transfer using colchicine-treated fibroblasts as donor cells. *Mol. Reprod. Dev.* 62, 300–306
78 Robl, J.M. *et al.* (2003) Artificial chromosome vectors and expression of complex proteins in transgenic animals. *Theriogenology* 59, 107–113
79 Hill, J.R. *et al.* (1999) Clinical and pathologic features of cloned transgenic calves and fetuses (13 case studies). *Theriogenology* 51, 1451–1465
80 Cibelli, J.B. *et al.* (2002) The health profile of cloned animals. *Nat. Biotechnol.* 20, 13–14
81 Pursel, V.G. *et al.* (1990) Expression and performance in transgenic pigs. *J. Reprod. Fertil. Suppl.* 40, 235–245

GENERAL AND COMPARATIVE ENDOCRINOLOGY 89, 51–61 (1993)

Recombinant Carp (*Cyprinus carpio*) Growth Hormone: Expression, Purification, and Determination of Biological Activity *in Vitro* and *in Vivo*

MIRA FINE,* EDNA SAKAL,* DORIT VASHDI,* VIOLET DANIEL,† AVIGDOR LEVANON,‡ ORLI LIPSHITZ,‡ AND ARIEH GERTLER*,[1]

*Department of Biochemistry and Human Nutrition, The Faculty of Agriculture, The Hebrew University of Jerusalem, Rehovot 76100, Israel; †Department of Biochemistry, The Weizmann Institute of Science, Rehovot 76100, Israel; and ‡Biotechnology General (Israel), Rehovot 76326, Israel

Accepted June 18, 1992

Carp growth hormone (cGH) cDNA (Koren *et al.*, 1989) was cloned under the control of λ-phage P_LO_L promoter and λcII ribosomal binding site into pBR322 plasmid to enable its expression in *Escherichia coli* A1645 that produces constitutively the thermolabile λ repressor cI857. Temperature shift to 42° abolished the repression, resulting in a high level of cGH expression. The bacterially expressed cGH protein, contained within the refractile body pellet, was solubilized in 4.5 *M* urea, refolded, and purified on Q-Sepharose column by stepwise elution with NaCl. The bioactive fraction was eluted at 0.2 *M* NaCl at a yield of 10–15%. This fraction contained predominantly (95%) 21.5-kDa monomeric cGH. The activity of cGH *in vitro* was bioassayed using Nb2-11C lymphoma cells (containing lactogenic receptors) and 3T3-F442A preadipocyte cells (containing somatogenic receptors). Bioactivity was found to be 0.01 and 6–10% that of human GH, respectively. *In vivo* cGH activity was measured by weekly ip injection in juvenille carp fed a low (23%) protein diet. Over a 6-week period, cGH increased the growth rate by 38% compared to fish injected with vehicle only. Identical injections with bovine GH yielded only a 21% increase.

Growth hormones (GHs) are 21- to 22-kDa proteins secreted by the anterior pituitary gland in vertebrates. Several GHs have been successfully cloned and expressed in *Escherichia coli* (Goeddel *et al.*, 1979; Seeburg *et al.*, 1983). Consequently, large-scale production by recombinant techniques has enabled their commercial use for pharmacological intervention. Human GH has been successfully used for the treatment of dwarfism in children (Jorgensen *et al.*, 1990) and for replacement therapy in adults (Christiansen and Jorgensen, 1991). Animal GHs have been used to promote growth and alter body composition in pigs (Etherthon *et al.*, 1987) and to increase milk production in cows (Bauman *et al.*, 1985). More recently several groups have reported the cloning of GHs from a variety of fish species (Sekine *et al.*, 1985; Agellon and Chen, 1986; Sato *et al.*, 1988), including carp (Koren *et al.*, 1989; Chao *et al.*, 1989). The use of GH as a pharmacological growth-promoting tool is of particular interest in fish, due to its potential to decrease dietary protein requirements (Hertz *et al.*, 1992). Fish, in contrast to mammals, require a 2- to 3-fold higher level of dietary protein to maintain their maximal growth rate. For most fish, the required dietary protein levels range from 35 to 55% (Wilson and Halver, 1986). Less than half of that is utilized for body protein anabolism, taking into account actual deposition of body protein and a normal biological value for consumed protein. The rest is used for energy (Pfeffer, 1982). We previously found that feeding a low (23%) pro-

[1] To whom correspondence should be addressed.

0016-6480/93 $4.00

tein diet for a period of 12 weeks to common carp resulted in an average weight gain which was 24% lower than that of fish fed a high (38%) protein diet. Weekly injections of recombinant bovine GH had no effect on fish fed the high protein diet, but increased the growth rate in fish fed low protein diets to the level achieved by the high protein diet (Hertz *et al.*, 1992). Since chronic use of an exogenous heterologous hormone could result in an immunogenic response, the present work describes the preparation, purification, characterization, and biological activity assessment of a recombinant carp GH (cGH).

MATERIALS AND METHODS

Materials

Recombinant bovine growth hormone (bGH) and human growth hormone (hGH) were obtained from Biotechnology General, Inc. (Israel). Human insulin-like growth factor 1 (IGF-I) was prepared though recombinant DNA technology (Saito *et al.*, 1986) and obtained, together with anti-IGF-I serum, from Fujisawa Pharmaceutical Co. (Osaka, Japan). Carrier-free $Na^{125}I$ was purchased from the New England Nuclear Co. (Boston, MA), and α-Methyl[^{3}H]thymidine (70–85 Ci/mmol) was purchased from Amersham (Buckinghamshire, England). Molecular weight markers for gel electrophoresis, Fisher's medium, RPMI 1640 medium, DME/F-12 medium, lysozyme, and BSA (RIA grade) were obtained from Sigma Chemical Co. (St. Louis, MO). Reagents for SDS–PAGE and a protein assay kit were purchased from Bio-Rad Laboratories (Richmond, CA). Lactogen-free horse serum and fetal calf serum (FCS) were obtained from Labotal Co. (Jerusalem, Israel). Superdex 75 column and Q-Sepharose (fast flow) were purchased from Pharmacia LKB Biotechnology AB (Uppsala, Sweden). All other chemicals were of analytical grade.

SDS–Polyacrylamide Gels

SDS–PAGE was carried out in 15% gels according to the method of Laemmli (1970). Gels were stained with Coomassie brilliant blue R.

Construction of cGH Expression Plasmid pO.L-2-FGN

The characterization of cGH cDNA isolated from a carp pituitary cDNA library in λ-gt11, has been described previously (Koren *et al.*, 1989). To enable the expression of cGH in a bacterial host, the plasmid pO.L-2-FGH was constructed (Fig. 1). To facilitate its construction the cGH fragment from plasmid p14-P-C was first subcloned into a pBR322-derived plasmid.

Plasmid pO.L-2-FGH is about 4560 bp long and consists of the following elements: a 270-bp fragment containing the λ phage P_LO_L regulatory region and the N gene utilization site (nut L); a 70-bp sequence containing the λcII ribosomal binding site (RBS); a 1021-bp fragment that includes the entire coding region of the mature cGH preceded by an ATG translation initiation site and terminating with TAA stop codon, followed by 3′ untranslated sequence; the ribosomal transcription termination sequences (T_1T_2); the Amp^R-β-lactamase gene of pBR322; and the latter's origin of replication.

Expression of cGH in E. coli

Expression plasmid pO.L-2-FGH was propagated in *E. coli* strain A1645, which constitutively produces the thermolabile λ repressor c1857 and the transcription antitermination N gene produce (Roberts, 1969). At 30°, the repressor binds to O_L and prevents transcription from the strong P_L promoter. When the temperature is raised to 42° repression is abolished, permitting transcription.

Analysis of Induced Culture

Cells were harvested by centrifugation and suspended in 50 m*M* potassium phosphate buffer (pH 7.8). Extracts were sonicated for 90 sec (3 × 30) in a W-375 sonicator (Heat Systems, Ultrasonics, Inc., USA) and partitioned to soluble and insoluble fractions by centrifugation at 10,000*g* for 5 min. The pellet was resuspended in the same volume of 50 m*M* potassium phosphate buffer (pH 7.8). Aliquots were lysed in 3× sample buffer with 30% (v/v) glycerol and analyzed on 15% SDS–PAGE gels.

Purification of cGH

Precipitated *E. coli* cells (80 g), which had been stored at −80°C were thawed, homogenized, and incubated for 30 min in 10 m*M* EDTA, pH 8.0, in the presence of 0.5 mg/ml lysozyme. The cells were sonicated and pelleted by centrifugation for 30 min at 25,000*g*. The pellet, containing the refractile bodies, was washed twice with distilled water by sonication, reprecipitated, and solubilized in 1.6 liter of 4.5 *M* urea buffered with 10 m*M* Tris–HCl, pH 8.0. After removal of the insoluble material, the concentration of Tris–HCl was raised to 40 m*M*, the pH was raised to 11.3, and cysteine was added to a concentration of 0.1 m*M*. After stirring for 48 hr at 4°, the solution was loaded

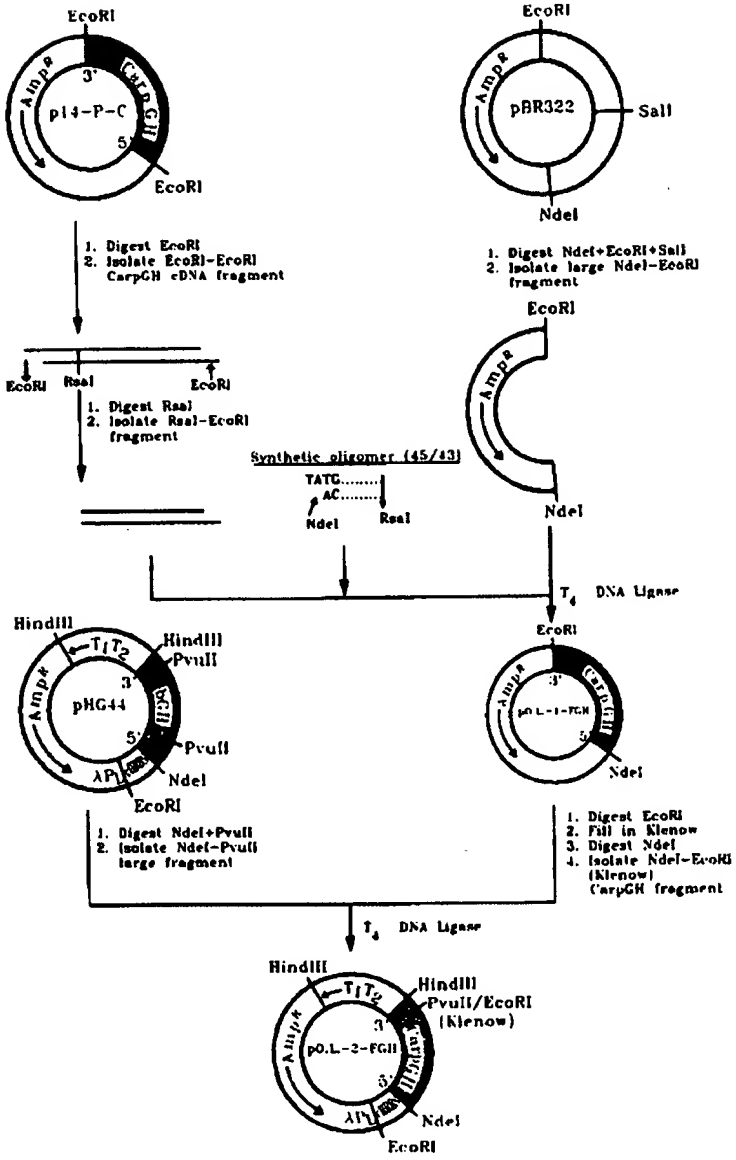


FIG. 1. Construction of cGH vector. Expression plasmid pO.L-2-FGH was obtained via manipulation of the cGH cDNA plasmid p14-P-C. Plasmid p14-P-C was digested with *Eco*R1 endonuclease. The *Eco*R1–*Eco*R1 (~1180 bp) fragment coding for the entire cGH cDNA was isolated and digested with *Rsa*1 endonuclease. The *Rsa*1–*Eco*R1 DNA fragment (~1020 bp) was isolated and the missing 5' end sequence of the mature cGH was reconstructed using a synthetic DNA fragment with the following sequence:

5'- TATGTCAGACAACCAGCGGCTCTTCAATAATGCAGTCATTCGTGT - 3'

3' - ACAGTCTGTTGGTCGCCGAGAAGTTATTACGTCAGTAAGCACA - 5'

Ligation was carried with the *Nde*1–*Eco*R1 isolated DNA fragment derived from plasmid pBR322. The newly obtained plasmid was designated pO.L-1-FGH. To facilitate expression under the control of bacteriophage λ regulatory sequences, plasmid pO.L.-1-FGH was digested with *Eco*R1 endonuclease and the sticky ends of the *Eco*R1 site were filled in with the large fragment of DNA polymerase (Klenow). This was followed by digestion with *Nde*1 endonuclease. The *Nde*1-blunted DNA fragment of the cGH was then subcloned into the bGH expressor plasmid pHG44 digested with *Pvu*11 and *Nde*1 endonucleases. The newly obtained plasmid was designated pO.L-2-FGH.

onto a Q-Sepharose column (1.5 × 16 cm) equilibrated with 20 m*M* Tris–HCl, pH 8.0, at 4°. Elution was carried out using a discontinuous NaCl gradient in the same buffer at a rate of 40 ml/hr, and 5-ml fractions were collected. Protein concentration was determined by absorbance at 280 nm and monomer content was determined by gel-filtration chromatography on a Superdex 75 column.

Determination of the Monomer Content

HPLC gel-filtration chromatography was performed on 200-μl aliquots of either Q-Sepharose column-eluted fractions or freeze-dried samples dissolved in H_2O (0.4–0.5 mg/ml), in a Merck–Hitachi apparatus equipped with a D-2000 integrator and an L-6200 controller. A Superdex (1 × 30 cm) column. preequilibrated with 0.05 *M* NaH_2PO_4 and 0.1 *M* $Na_2(SO)_4$, at pH 7.3 was used. The column was developed in the same buffer at a rate of 1.0 ml/min at room temperature. Protein content was monitored by absorbance at 280 nm and quantified using the D-2000 integrator. The retention time of several known proteins, such as bovine serum albumin, hGH, and α-lactalbumin was determined to calibrate the column. The molar extinction coefficient at 280 nm was calculated using the molar extinction coefficients of tyrosine (1197) and tryptophan (5559).

Determination of the Amino Terminal Sequence

Automated Edman degradation technique was used to determine the amino-terminal protein sequence. Degradations were performed on an ABI Model 470A gas phase sequencer (Foster City, CA), according to Hunkapiller *et al.* (1983), using the standard sequencing cycle. The respective PTH–amino acid derivatives were identified by RP-HPLC analyses in an on-line fashion, using an ABI Model 120A PTH analyzer fitted with a Brownlee 2.1-mm-i.d. PTH-C_{18} column.

Radioimmunoassay of IGF-I

IGF-I was measured by radioimmunoassay using human recombinant IGF-I as the tracer. IGF-I was iodinated by the chloramine-T method according to a previously described protocol (Gertler *et al.*, 1984). Serum samples (250 μl) were chromatographed on an octadecasil–silica column according to Funkenstein *et al.* (1989) to remove the IGF-I binding proteins. The radioimmunoassay and the separation of free and antibody-bound ^{125}I-labeled IGF-I were carried out by the PEG method (Gluckman *et al.*, 1983). The detectable range of the assay was 10–2500 pg IGF-I/tube.

Nb_2-11C Lymphoma Cell Culture

Nb_2-11C lymphoma cell-line culture, synchronization of Nb_2-11C cells in the G_0/G_1 phase, and monitoring of cell proliferation were carried out as described previously (Gertler *et al.*, 1985). This bioassay is highly specific for lactogenic hormones such as hGH or prolactins (Tanaka *et al.*, 1980).

3T3-F442A Preadipocyte Bioassay

The present bioassay uses 3T3-F442A rat preadipocytes as a model. These cells possess a well-defined somatogenic receptor that recognizes not only hGH, but other GHs as well (Nixon and Green, 1983; Uchida *et al.*, 1989). It has been reported that hGH is required for differentiation of these cells to adipocytes (Hayashi *et al.*, 1981). 3T3-F442A preadipocytes were seeded at a level of 22000 cells/well in a 24-well plate containing DMEM medium with 10% FCS. After 24 hr the cells were transferred to a serum-free medium containing supplemental proteins (Corin *et al.*, 1990) and increased concentrations of cGH or hGH and cultured for 5 days. The medium was then changed to DMEM containing 4% FCS and DNA synthesis was determined 20 hr later following a 2-hr pulse of 1 μCi/ml [^{3}H]thymidine. Human GH inhibits DNA synthesis in a dose-dependent manner with IC_{50} = 0.1 ng/ml (Vashdi *et al.*, 1991).

Experimental Animals

Common carp (*Cyprinus carpio*) was used as the experimental animal. Maintenance and diets were as described previously (Hertz *et al.*, 1989a,b). Growth experiments were performed in large tanks containing 18–20 fish each (one tank per treatment), for a period of 6 weeks. Injections of cGH or bGH dissolved in 0.05% $NaHCO_3$ to a concentration of 1 mg/ml (5 μg/g body wt) and 0.05% $NaHCO_3$ (vehicle) were carried out once a week at which time fish were also individually weighed. Twenty-four hours after the last injection blood samples were withdrawn for IGF-I determination. Since the absolute amount of blood was very small, samples from three or four fish given the same hormonal treatment were randomly pooled.

Statistical Analysis

One-way analyses of variance were performed. When the *F* values were significant ($P < 0.05$), means of the various treatments were compared using Duncan's multiple range test. All parametric data are expressed as the mean ± SEM.

RESULTS

Expression of cGH in E. coli

The production of cGH from plasmid pO.L-2-FGH after induction of *E. coli* A1645 culture by temperature shift to 42° is shown in Fig. 2. Carp GH accumulated up to 120 min, at which stage the recombinant protein constituted about 28% of total cellular protein, as determined by densitometric scanning of the Coomassie-stained gel. Figure 2 also indicates that the recombinant cGH protein is found only in the nonsoluble refractile bodies (lanes 5 and 7) and is absent in the cell extracts supernatants (lanes 4 and 6).

Purification and Characterization of Bacterial Expressed cGH

The cGH within the refractile body pellet was solubilized in 4.5 *M* urea, refolded at pH 11.3 in the presence of catalytic amounts of cysteine, and purified on a Q-Sepharose column, as shown in Fig. 3. The fraction eluted at 200 m*M* NaCl produced a nonsymmetrical right-skewed peak, as determined by absorbance at 280 nm. Fractions corresponding to the symmetric part of the peak were collected, pooled, dialyzed against 0.05% $NaHCO_3$, and freeze-dried. Yield varied between 10 and 15 mg protein per 80 g of precipitated *E. coli* cells. A large amount of material absorbing at 280 nm was eluted at 1 *M* NaCl (not shown). This material was composed mainly of nucleic acids, as suggested by the high 260/280 nm ratio. The Q-Sepharose column fractions were analyzed by gel filtration on Superdex column (Fig. 4). The fraction eluted at 200 m*M* NaCl (Fig. 3), consisted of over 95% monomeric cGH (Fig. 4B), having retention time equal to that of recombinant hGH. Analysis of the pooled 300 m*M* NaCl eluates revealed drastically reduced monomer content, with the

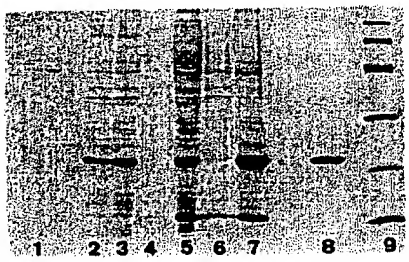


FIG. 2. Accumulation of cGH in *E. coli* strain A1645. Cells were grown at 30° in L broth containing ampicillin (100 μg/ml). Induction of exponentially growing cultures was performed by a temperature shift to 42°. Extracts of samples taken at Time 0 and 60 and 120 min postinduction (lanes 1–3, respectively) were electrophoretically analyzed on a 15% SDS–PAGE under reducing conditions. Each lane contained the equivalent of 50 μl of culture at A_{600} = 1. Lane 4, soluble protein fraction in the supernatant of culture, 60 min postinduction at 42°; lane 5, nonsoluble pellet of culture, 60 min postinduction; lane 6, soluble protein fraction in the supernatant of culture, 120 min postinduction at 42°; lane 7, nonsoluble pellet of culture, 120 min postinduction; lane 8, 2 μg of purified recombinant hGH; lane 9, protein molecular weight markers (from top to bottom in kDa): 94, 67, 43, 30, 20.1, and 14.3 Gel was stained with Coomassie brilliant blue.

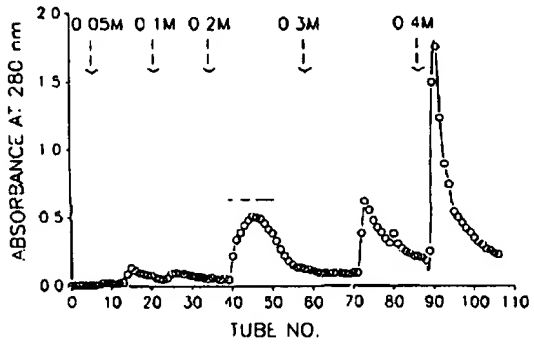


FIG. 3. Purification of proteins extracted from refractile bodies on a Q-Sepharose column. The column (1.5 × 16 cm) was equilibrated with 20 m*M* Tris–HCl, pH 8.0, at 4°. The fraction containing refractile bodies' proteins solubilized in 4.5 *M* urea in 40 m*M* Tris–HCl at pH 11.3 (1500 ml) was applied to the column at a rate of 40 ml/hr. Subsequently the column was washed with 40 ml 20 m*M* Tris–HCl, pH 8.0. The eluate was not collected. Then elution was carried out using a discontinuous NaCl gradient in the same buffer at 40 ml/hr and 5-ml fractions were collected. The protein concentration was determined by absorbance at 280 nm. The bar over the peak eluted with 0.2 *M* NaCl was taken for dialysis and lyophilization.

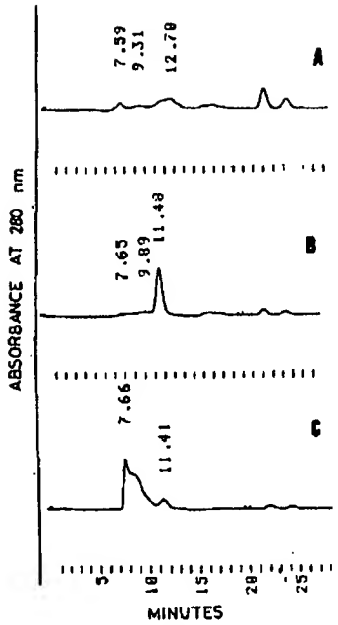


FIG. 4. Determination of the monomer content of cGH fractions eluted from the Q-Sepharose column, by HPLC gel chromatography on a Superdex 75 (1 × 30 cm) column. The protein content was monitored by absorbance at 280 nm and quantified using a Merck-Hitachi D-2000 integrator and the numbers within the figure indicate retention times. A, B, and C are fractions eluted with 100, 200, and 300 m*M* NaCl, respectively. For more details see text.

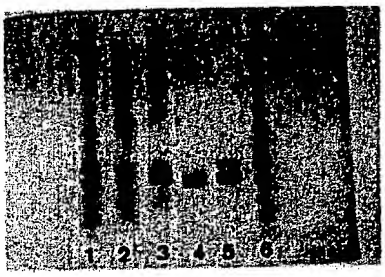


FIG. 5. Electrophoretic analysis of cGH at different stages of purification. Aliquots were analyzed by electrophoresis on a 15% SDS–PAGE under nonreducing conditions. Lanes 1 and 6, protein molecular weight markers (from top to bottom in kDa): 66, 34.7, 24.0, 18.4, and 14.3; lane 2, nonsoluble pellet of culture 120 min postinduction; lane 3, refractile bodies, prior to refolding; lane 4, Q-Sepharose column-purified cGH; lane 5, 5 μg of purified recombinant hGH. Gel was stained with Coomassie brilliant blue.

main species consisting of oligomeric and dimeric forms (Fig. 4C). The fraction eluted at 400 m*M* NaCl had a similar profile (not shown), but hardly any 21- to 22-kDa protein could be detected in the 100 m*M* NaCl eluate (Fig. 4A). The protein profile obtained by SDS–PAGE under denaturating, nonreducing conditions for different stages of purification is presented in Fig. 5. The expressed cGH accounted for up to 28% of the total protein in the whole cell extract and up to 70% in the refractile bodies. Analysis of the fraction eluted at 200 m*M* NaCl revealed over 95% purity and a MW of 21.5 kDa (compared to the theoretical value of 21.4 kDa calculated from the expected amino acid content—Koren *et al.*, 1989). The cGH bands from the cell extract and refractile bodies exhibited slightly slower mobility, probably because these cGH molecules had not yet refolded. Analysis of the purified cGH under reducing conditions (not shown) also revealed slower mobility indirectly implying a refolded state of the purified protein. Amino-terminal analysis of five amino acids resulted in an expected sequence (Koren *et al.*, 1989), namely, Ser–Asp–Asn–Glu–Arg, with respective yields of 446, 624, 735, 718, and 565 pmol/cycle. A considerable amount of Met (90 pmol) was found in the first cycle. Since the recovery of Ser is usually low, the yields of cycles 2–4 were used to calculate the Met-cGH percentage. Met-cGH was found to account for not more than 12–13%, indicating that most of the Met had been processed in the *E. coli* cells after expression.

Biological Activity of the Purified cGH

The biological activity of the purified cGH was assessed by two *in vitro* and one *in vivo* bioassays. In the somatogenic receptor-mediated bioassay, in 3T3-F442A preadipocytes, 5 days of exposure to 25 ng/ml cGH resulted in ~70% inhibition of FCS-stimulated [^{3}H]thymidine incorporation (Fig. 6A). This effect was similar to the maximum effect obtained with hGH, which served as a reference. In order to achieve

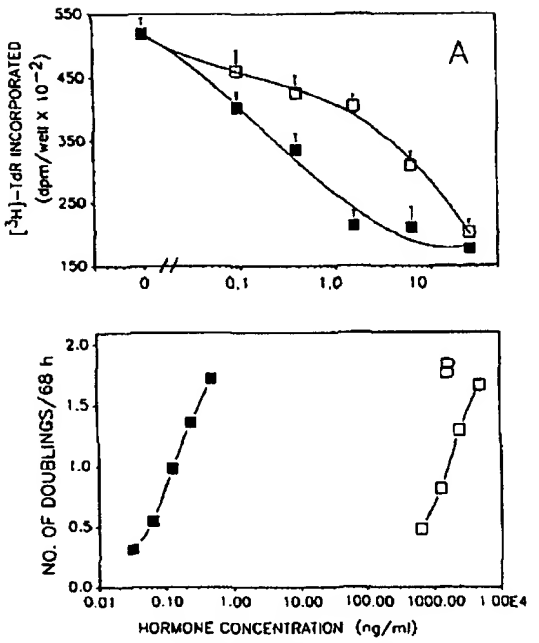


FIG. 6. (A) The anti-mitogenic effect of cGH (□) and hGH (■) in 3T3-F442A preadipocytes. After seeding, the cells were transferred to a serum-free medium containing supplemental proteins (Corin *et al.*, 1990) and increasing concentrations of cGH or hGH and cultured for 5 days. Then the medium was changed to one containing 4% FCS, and DNA synthesis was determined 20 hr later by a 2-hr pulse of 1 μCi/ml [^{3}H]thymidine. (B) The effect of cGH (□) and hGH (■) on the proliferation rate of Nb_2-11C lymphoma cells. Doubling rates were calculated after 68 hr from log (no of cells in the presence of the hormone/no. of cells in the absence of the hormone) log 2.

the half-maximum effect, however, approximately 15-fold more cGH was required, probably due to its relatively low affinity for somatogen receptors in these cells. In the Nb_2-11C lymphoma cell bioassay, mediated through lactogen-type receptors, hGH or prolactins can stimulate proliferation with a half-maximum effect at 0.1–0.2 ng/ml (Tanaka *et al.*, 1980). Carp GH was also capable of stimulating the proliferation of Nb_2-11C lymphoma cells, although a 10^4-fold higher concentration was required to the achieve the half-maximum effect (Fig. 6B). Fractions eluted from the Q-Sepharose column at 100, 300, and 400 m*M* NaCl (Fig. 3) were inactive in both bioassays.

Further assessment of the purified cGH's biological activity was performed *in vivo* by determining its growth-promoting effect. The experiment was carried out for 6 weeks using juvenile carps fed low protein (23%) diets. This protein level was chosen in view of former results showing the growth-promoting activity of bGH, when administered along with low but not high protein diets (Hertz *et al.*, 1992). Figure 7 shows that as early as 2 weeks after the first weekly injection, both cGH and bGH had caused a significantly ($P < 0.05$) higher weight gain than that in fish injected with vehicle. Similar results were obtained through the fourth week. In the fifth and sixth weeks, however, the growth-promoting effect of bGH was attenuated compared to that of cGH. Overall weight gain after 6 weeks in fish treated with bGH was 21% higher than that in fish injected with vehicle, while injections with cGH yielded a 38% increase.

The day after the last injection the fish were bled and the level of IGF-I in the sera was determined using heterologous radioimmunoassay. The respective values for fish treated with bGH and cGH (in ng/ml) were (mean ± SEM) 1.00 ± 0.21 and 0.93 ±

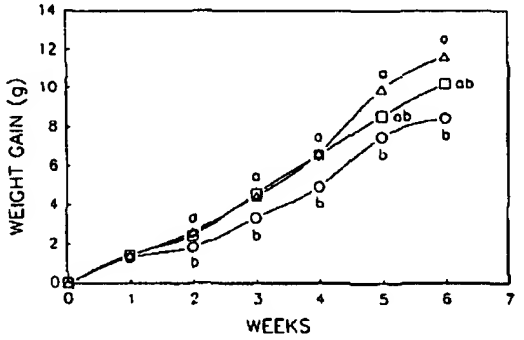


FIG. 7. The effect of cGH (△), bGH (□), or vehicle (○) on weight gain in carps fed a low protein (23%) diet. Intraperitoneal injections of the hormones dissolved in 0.05% $NaCO_3$ at 1 mg/ml (5 μg/g body wt) were carried out once a week, at which time fish were also individually weighed. Mean results are presented (n = 19–20). Means not designated with the same letter are significantly different ($P < 0.05$).

0.13. Both values were significantly ($P < 0.05$) higher than those in fish injected with vehicle (0.47 ± 0.06). These values are, however, 200- to 500-fold lower compared to mammals and 10-fold lower compared to trout (Daughaday *et al.*, 1985). They were also ~200-fold lower than the only previously reported value for carp (Funkenstein *et al.*, 1989), who used antiserum raised against the synthetic 53–70 IGF-I fragment. Thus the IGF-I values reported by us may not be accurate, because in the present work an anti-human IGF-I antiserum was used. There is no current information concerning the immunological cross-reactivity of human and carp IGF-I.

DISCUSSION

Heat-stimulated *E. coli* cells transformed with the pO.L-2-FGH vector containing cGH cDNA, expressed high levels of cGH, which accounts 28% of total cell proteins. Expressed cGH was identified as the main protein (>70%) in the refractile bodies prior to solubilization in urea and refolding (Figs. 2 and 5). Further purification on a Q-Sepharose column yielded a monomeric 21.5-kDa protein, which was 95% homogeneous, as demonstrated by SDS–PAGE (Fig. 5). The electrophoretic mobility of this purified cGH under nonreducing conditions was higher than that of the corresponding cGH band in whole cell or refractile body extracts, indicating refolding. The identity of the purified cGH was further verified by amino-terminal sequence analysis.

The monomeric form of cGH, eluted from the Q-Sepharose column at 200 m*M* NaCl, was the only fraction that exhibited biological activity both *in vitro* and *in vivo*. Although the overall somatogenic receptor-mediated anti-mitogenic effect of cGH in 3T3-F442A preadipocytes was similar to that of hGH, the IC_{50} value was 10- to 15-fold higher, raising the question of whether this lower activity is due to improper refolding or results from an intrinsically lower affinity of cGH for this receptor. The observation that the *in vivo* growth-promoting activity of cGH exceeded that of bGH argues in favor of the second possibility. It should be noted that the biological activity of bGH in the 3T3-F442A preadipocyte bioassay is almost equal to that of hGH (Vashdi *et al.*, 1991). This conclusion is also consistent with binding studies of radiolabeled cGH to carp liver microsomal fraction. Competition studies with cGH, hGH, or bGH indicated that the last two hormones have at least 50-fold lower affinity toward cGH receptors, (Fine and Gertler, unpublished). Purified cGH was also active in the Nb_2-11C lymphoma cell bioassay, though up to 10^4-fold higher concentrations were required to achieve a biological effect equal to that of hGH. This assay, which is mediated through lactogenic receptors, was assumed to be nonresponsive to GHs other than those or primates (Tanaka *et al.*, 1980). With the advent of recombinant GHs that can be tested for intrinsic activity in micromolar concentrations, uncontaminated by other pituitary hormones such as prolactin, this assumption should be revised. In point of fact, recombinant bGH was also found to be active in this assay, although its activity was ~5-fold lower than that of cGH (Cohen, Sakal, and Gertler, unpublished).

Carp GH exhibited significant growth-promoting activity when injected into juvenile fish fed low protein diets, similar to that formerly reported for bGH (Hertz *et al.*, 1992). In the first 4 weeks of the experiment, bGH and cGH were equally active. After 6 weeks, however, cGH was more active than bGH, though the difference was statistically borderline. The reason for the attenuation of bGH's effect may be related to a possible immunogenic effect, although direct evidence of this is lacking. These results parallel the growth-promoting effects of other fish GHs demonstrated in trout (Sato *et al.*, 1989), coho salmon (McLean *et al.*, 1990) and eel (Sugimoto *et al.*, 1990).

Most of the protein present in the refractile bodies was eluted from the Q-Sepharose column at 300 and 400 m*M* NaCl. Although SDS–PAGE under reducing conditions revealed the 22-kDa cGH species to be the main protein (not shown), gel filtration under nondenaturing, nonreducing conditions showed the fractions to be composed of a mixture of high molecular weight oligomeric forms (Fig. 4). These fractions were also devoid of the *in vitro* biological activity exhibited by the monomeric hormone, indicating probably an improper refolding. Other recombinant hormones such as hGH, bovine placental lactogen, and their analogues refold properly under similar conditions, leading to high yields of the monomeric form (Binder *et al.*, 1999; Gertler *et al.*, 1992). The low yield of a properly refolded monomeric fraction in cGH may result from the fact that, in contrast to other mammalian and nonmammalian GHs (Nicoll *et al.*, 1986; Kimura, 1991), cGH has five rather than four Cys residues. Four residues, located at positions 48, 161, 178, and 186 (Koren *et al.*, 1989; Chao *et al.*, 1989) are homologous to those of other GHs, while the fifth residue is located at position 123. In other GHs this position is occupied by noncharged amino acids. It is well established that an unpaired number of cysteine residues has an adverse effect on proper refolding, leading to the formation of oligomers (Buchner and Rainer, 1991). Thus mutation of Cys_{123} to a noncharged but similar-sized amino acid such as serine or alanine may lead to the creation of a mutant which should retain full biological activity and give a higher yield of the monomeric form. Such a mutant, having only four Cys residues, would be advantageous for possible practical uses in the carp industry and is now being prepared. Our former results, showing that the addition of bGH to low protein diets may compensate for the lower growth rate observed in carp (Hertz *et al.*, 1992), support the use of cGH, once an efficient delivery system has been developed, as being not only economically feasible but also preferable, due to the lower potential contamination of the environment by nitrogenous compounds.

ACKNOWLEDGMENTS

The authors thank Dr. A. Nimrod from the Biotechnology General (Israel) for the recombinant bGH and hGH, Ms. N. Chapnik-Cohen for technical help in the Nb_2 bioassays, and Ms. Marsha Rockitter for typing the manuscript. This research was supported by Grant AQ5/IVGR537 from the joint German–Israeli research projects: the Bundesministerium fuer Forschung und Technologie (Germany) and the Ministry of Science and Technology (Israel).

REFERENCES

Agellon, L. B., and Chen, T. T. (1986). Rainbow trout growth hormone: Molecular cloning of cDNA and expression in *E. coli*. *DNA* 5, 463–471.

Bauman, D. E., Eppard, P. J., DeGeeter, M. J., and Lanza, G. M. (1985). Responses of high producing dairy cows to long-term treatment with pituitary somatotropin and recombinant somatotropin. *J. Dairy Sci.* 68, 1352–1362.

Binder, L., Gertler, A., Elberg, G., Guy, R., and Vogel, T. (1990). Site-directed mutations of human growth hormone that selectively modify its lactogenic activity and binding properties. *Mol. Endocrinol.* 4, 1060–1068.

Buchner, J., and Rainer, R. (1991). Renaturation, purification, and characterization of recombinant Fab-fragment produced in *Escherichia coli*. *Biotechnology* 9, 157–162.

Chao, S. C., Pan, F. M., and Chang, W. C. (1989). Purification of carp growth hormone and cloning of the complementary DNA. *Biochim. Biophys. Acta* 1007, 233–236.

Christiansen, J. S., and Jorgensen, J. O. L. (1991). Beneficial effects of GH replacement therapy in adults. *Acta Endocrinol. (Copenhagen)* 125, 7–13.

Corin, R. E., Guller, S., Wu, K. Y., and Sonenberg, M. (1990). Growth hormone and adipose differentiation: Growth hormone-induced antimitogenic state in 3T3-F442A preadipose cells. *Proc. Natl. Acad. Sci. USA* 87, 7507–7511.

Daughaday, W. H., Kapadia, M., Yanow, C. E., Fabrick, K., and Mariz, I. K. (1985). Insulin-like growth factors I and II of nonmammalian sera. *Gen. Comp. Endocrinol.* 59, 316–325.

Etherthon, T. D., Wiggins, J. P., Evock, C. M., Chung, C. S., Rebhun, J. S., Walton, P. E., and Steele, N. C. (1987). Stimulation of pig growth performance by porcine growth hormone: Deter-

mination of the dose-response relationship. *J. Anim. Sci.* 64, 433–443.

Funkenstein, B., Silbergeld, A., Cavari, B., and Laron, Z. (1989). Growth hormone increases plasma levels of insulin-like growth factor (IGF-1) in a teleost, the gilthead seabream (*Sparus aurata*). *J. Endocrinol.* 120, R19–21.

Gertler, A., Ashkenazi, A., and Madar, Z. (1984). Binding sites of human growth hormone and ovine and bovine prolactins in the mammary gland and the liver of lactating dairy cow. *Mol. Cell. Endocrinol.* 34, 51–57.

Gertler, A., Walker, A., and Friesen, H. G. (1985). Enhancement of human growth hormone-stimulated mitogenesis of Nb_2 node lymphoma cells by 12-*O*-tetradecanoyl-phorbol-13-acetate. *Endocrinology* 116, 1636–1644.

Gertler, A., Hausner, S. D., Sakal, E., Vashdi, D., Staten, N., Freeman, J. J., and Krivi, G. G. (1992). Preparation, purification, and determination of the biological activities of twelve N-terminus truncated recombinant analogues of bovine placental lactogen. *J. Biol. Chem.* 267, 12655–12659.

Gluckman, P. D., Johnson-Barret, J. J., Butler, J. H., Edgar, B. W., and Gunn, T. R. (1983). Studies of insulin-like growth factor-I and -II by specific radioligand assays in umbilical cord blood. *Clin. Endocrinol.* 49, 405–413.

Goeddel, D. V., Heyneker, H. L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D. G., Ross, M. J., Miozzari, G., Crea, R., and Seeburg, P. H. (1979). Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. *Nature* 281, 544–548.

Hayashi, I., Nixon, T., Morikawa, M., and Green, H. (1981). Adipogenic and antiadipogenic factors in the pituitary and other organs. *Proc. Natl. Acad. Sci. USA* 78, 3969–3972.

Hertz, Y., Madar, Z., Hepher, B., and Gertler, A. (1989a). Glucose metabolism in the common carp (*Cyprinus carpio* L.): The effects of cobalt and chromium. *Aquaculture* 76, 255–267.

Hertz, Y., Epstein, N., Abraham, M., Madar, Z., Hepher, B., and Gertler, A. (1989b). Effects of metformin on plasma insulin, glucose metabolism, and protein synthesis in the common carp (*Cyprinus carpio* L.). *Aquaculture* 80, 175–187.

Hertz, Y., Fine, M., Madar, Z., Tchelet, A., Shechter, Y., and Gertler, A. (1992). The effect of different levels of dietary protein on metabolic/endocrine profile and growth of carp (*Cyprinus carpio L.*): Involvement of growth hormone and insulin. *In* "Progress in Aquaculture Research" (B. Moaz, V. Hilage, and H. Rosenthal, Eds.), European Aquaculture Society, special publication No. 13. Bredene, Belgium, in press.

Hunkapiller, M. W., Hewick, R. M., Dreyer, W. J., and Hood, L. E. (1983). High sensitivity sequencing with gas phase sequentor. *Methods Enzymol.* 91, 399–413.

Jorgensen, J. O. L., Moller, J., Moller, N., Lauritzen, T., and Christiansen, J. S. (1990). Pharmacological aspects of growth hormone replacement therapy: Route, frequency, and timing of administration. *Horm. Res.* 33, 77–82.

Kimura, A. (1991). Microbial production of fish (tuna) growth hormone. *Crit. Rev. Biotechnol.* 11, 113–127.

Koren, Y., Sarid, S., Ber, R., and Daniel, D. (1989). Carp growth hormone: Molecular cloning and sequencing of cDNA. *Gene* 77, 309–315.

Laemmli, U. K. (1970). Cleavage of structure proteins during assembly of the heat of bacteriophage. *Nature* 227, 680–685.

McLean, E., Donaldson, E. M., Dye, H. M., and Souza, L. M. (1990). Growth acceleration of coho salmon (*Oncorhynchus kisutch*) following oral administration of recombinant bovine somatotropin. *Aquaculture* 91, 197–203.

Nicoll, C. S., Mayer, G. L., and Russell, S. H. (1986). Structural features of prolactins and growth hormones that can be related to their biological properties. *Endocr. Rev.* 7, 169–203.

Nixon, T., and Green, H. (1983). Properties of growth hormone receptors in relation to the adipose conversion of 3T3 cells. *J. Cell. Physiol.* 115, 291–296.

Pfeffer, E. (1982). Utilization of dietary protein by salmoid fish. *Comp. Biochem. Physiol. B.* 73, 51–57.

Roberts, J. W. (1969). Termination factor for RNA synthesis. *Nature* 224, 1168–1174.

Saito, Y., Yamada, H., Niwa, M., and Ueda, I. (1986). Production and isolation of recombinant somatomedin C. *J. Biochem.* (*Tokyo*) 101, 123–134.

Sato, N., Watanbe, K., Murata, K., Sakaguchi, M., Kariya, Y., Kimura, S., Nonaka, M., and Kimura, A. (1988). Molecular cloning and nucelotide sequence of tuna growth hormone cDNA. *Biochim. Biophys. Acta* 949, 35–42.

Sato, N., Kawazoe, I., Tamai, T., Inogue, Y., Murata, K., Kimura, S., Nonaka, M., and Kimura, A. (1989). Purification and characterization of a recombinant tuna growth hormone produced in *Escherichia coli*. *J. Ferment. Bioeng.* 68, 79–83.

Seeburg, P. H., Sias, S., Adelman, J., DeBoer, H. A., Hayflock, J., Jhurani, P., Goeddel, D. V., and Heyneker, H. L. (1983). Efficient bacgerial expression of bovine and porcine growth hormones. *DNA* 2, 37–45.

Sekine, S., Muzukami, T., Nishi, T., Kuwana, Y., Saito, A., Sato, M., Itoh, S., and Kawauchi, H.

(1985). Cloning and expression of cDNA for salmon growth hormone in *Escherichia coli*. *Proc. Natl. Acad. Sci USA* 82, 4306–4310.

Sugimoto, S., Yamaguchi, K., and Yokoo, Y. (1990). Isolation and characterization of recombinant eel growth hormone expressed in *Escherichia coli*. *J. Chromatogr.* 515, 483–494.

Tanaka, T., Shiu, R. P. C., Gout, P. W., Beer, C. T., Noble, R. L., and Friesen, H. G. (1980). A new sensitive and specific bioassay for lactogenic hormones: Measurement of prolactin and growth hormone in human serum. *J. Clin. Endocrinol. Metab.* 51, 1058–1063.

Uchida, E., Hayakawa, T., Niimi, S., Tanaka, A., and Morikawa, M. (1989). Structural characteristics of growth hormone receptors on mouse 3T3 cells having different biological responses to growth hormone. *J. Mol. Endocrinol.* 3, 239–245.

Vashdi, D., Sakal, E., Elberg, G., and Gertler, A. (1991). Characterization of the receptor and the biological activity of bovine placental lactogen. *Proc. Annu. Meet. Israeli Endoc. Soc.*, Abstract 43.

Wilson, R. P., and Halver, J. G. (1986). Protein and amino acid requirement in fish. *Annu. Rev. Nutr.* 6, 225–244.

Antibodies produced in plants

A. Hiatt

Transgenic plant systems for the expression of mammalian antibodies offer opportunities for the study of plant metabolism and development. Agricultural production could provide virtually unlimited quantities of any antibody.

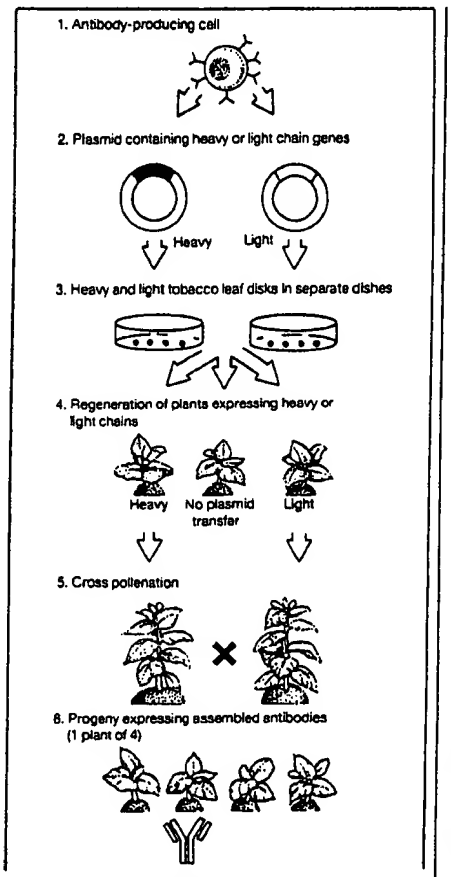
HETEROLOGOUS systems for the expression of mammalian antibodies will undoubtedly contribute a great deal towards our ability to isolate and manipulate immunoglobulins[1]. The latest heterologous host system for antibody synthesis is plants. Techniques to generate transgenic plants have been perfected to the point where a foreign protein can be targeted to an organ of choice as well as to subcellular compartments.

Production of antibodies by plant cells offers a variety of new possibilities for basic research in plant biology as well as for large-scale production of antibodies for use as therapeutic, diagnostic or affinity reagents. The unparalleled capacity and flexibility of agricultural production suggests that antibodies derived from plants may be significantly less expensive than antibodies from any other source. Moreover, antibodies in plants may become useful reagents for manipulating agronomic traits and possibly for ameliorating symptoms of pathogenic infections, as well as for isolating and processing environmental contaminants or industrial by-products.

Plant transformation

Successful expression of an antibody in tobacco has recently been reported[2]. A catalytic antibody[3] was chosen to test the ability of the tobacco cell to assemble and process immunoglobulin chains without compromising functionality. cDNAs encoding heavy and light chains were first inserted into *Agrobacterium tumefaciens*, a soil bacterium that has proven to be very useful for transforming many types of plant cells[4]. The *Agrobacterium* is responsible for transferring the DNA into the plant cell where it is subsequently integrated into the genome. Transformed plant cells are then regenerated to become mature plants[5].

The strategy used for antibody production was to transform tobacco leaf discs and regenerate separate plants expressing either the light or heavy chains (see figure). These plants were then sexually crossed to produce progeny-expressing functional antibody. Although the levels of expression varied widely, greater than one per cent of total protein constituted functional antibody in some plants. There is reason to believe that this level of expression can be augmented by using promoter elements capable of higher levels of transcription[6]. The antibody can easily be purified from homogenized leaves in one affinity purification step. The catalytic properties of the tobacco-produced antibody allow a precise evaluation of kinetic parameters such as K_m, K_i and K_{cat}; by these functional criteria, it is identical to the same antibody derived from hybridoma cells. Further characterization (for example, site of synthesis, secretion, glycosylation) will be reported elsewhere.



Production of antibodies in tobacco plants. Primary regenerants transformed with *Agrobacterium* containing heavy or light chain cDNAs are sexually crossed to enable assembly of a functional antibody in resulting progeny.

Of critical importance, is an evaluation of the immunogenicity of plant-derived antibodies in mammals. As plants do not contain sialyl transferase activity[7], the terminal residues of the carbohydrate on the heavy chain will be different from mammals. In all probability, they will consist of xylose, fucose, and/or *N*-acetylglucosamine[7]. The extent to which alterations in carbohydrate composition affect the biodistribution and serum clearance of the antibody remains to be determined.

Agricultural-scale production

Clearly, if antibodies are to be used for therapeutic purposes, techniques for large-scale production have to be developed. The high capacity and flexibility of agricultural production offers several advantages for obtaining antibodies: genetically stable seed stocks of antibody-producing plants can be isolated and stored indefinitely at low cost and the seed stock can be converted into a harvest of any quantity of antibody within one growing season.

Although tobacco has been used as the principle research tool to initiate the study of antibodies in plants, there may be more appropriate plants for production. A variety of common crop plants can be used as the production host. Acreages of perennial forage crops could be generated by clonal propagation or from seed and harvested numerous times in a growing season. The choice of species may depend on the quantity and nature of contaminants encountered during purification. Some candidates are alfalfa, soybean, tomato and potato.

As large-scale production of antibodies is not yet commonplace, appropriate techniques for the purification of hundreds or thousands of grams have yet to be perfected. The cost of agriculturally-produced antibodies is likely to be considerably less than antibodies produced from hybridoma cells or ascites fluid. For example, if antibodies were expressed in soybean and constituted one per cent of total protein in soybean meal, a kilogram of antibody could, hypothetically, be produced for less than $100 (US). This extrapolation is based on current costs for soybean production and does not take into account numerous hidden costs such as the cost of development and propagation of a sufficiently large and genetically characterized seed stock. In addition, the efficiency with which antibodies can be produced in specialized organs such as seeds or fruit is still not known.

Growth regulation

Plant growth and development is controlled by a limited number of low molecular weight hormones such as indoleacetic acid, ethylene, benzylaminopurine and a variety of more complex organic

molecules[8]. Little is known about the biosynthetic pathways or the mechanism of action of these hormones. However, by expression within the plant cell of monoclonal antibodies that recognise these hormones, it may be possible to evaluate developmental and metabolic events that are controlled by their free titre. Ideally, one would want to control the expression of the antibody as well as target expression to different organs or subcellular locations. In this way, activities of the hormone at various developmental stages could be unravelled.

Pathogen resistance

Antibodies against hormones are just one area where expression of an endogenous antibody could aid plant research. Another example is infection of plants by pathogens. Although many fungal, bacterial and viral pathogens have been characterized with respect to the genetics of host-pathogen interactions, very few have been thoroughly investigated at the biochemical level. In some instances, however, pathogen-related proteins or other organic molecules have been shown to be necessary for pathogenesis[9,10].

Expression of an intracellular antibody that binds antigens essential for pathogenesis may ameliorate the symptoms of the infection by reducing the functional titre. The advantage of this strategy is twofold: first, it would not require isolation of genes involved in synthesis of the target antigen (as with anti-sense RNA expression); and second, pools of antigen which may be localized in subcellular compartments can be the specific target, leaving other pools unaffected. Clearly, the success of this approach will depend on a much more detailed understanding of the behaviour of antibodies in plants. Whereas antibodies have been successfully expressed intracellularly in both yeast and mammalian cells[11,12], attempts to assemble immunoglobulin chains in the cytosol of plants have been unsuccessful.

Current efforts are focusing on alternative methods which would by-pass the requirement for assembly of two immunoglobulin chains (for example, single chain antigen-binding constructs)[13]. In addition, attempts to localize an antigen-binding capacity to chloroplast and vacuole are in progress. Once we have a clear picture of the assembly, stability and functionality of targeted immunoglobulins, appropriate strategies for localized antigen binding can be devised.

Biofiltration

One of the key differences between plant cells and those of other organisms is the structure and characteristics of the surrounding cell wall. The mechanical strength and contiguous nature of plant cell walls is largely responsible for the rigidity of the entire plant. The diameter of pores in the cell wall imposes a restriction on the size of molecules that are freely permeable. This exclusion limit lies between 35 and 50 Å and corresponds to a molecular weight of less than 20,000 for a globular protein. Clearly, antibodies are too large to be freely permeable[14]. Consequently, expression of an antibody in a plant cell is equivalent to producing a binding and retention capacity within a semipermeable membrane. Any antigen with a molecular weight of less than 20,000 (for example, environmental pollutants, industrial by-products, pesticides and herbicides) might be collected and retained by a plant expressing an antibody that is functional *in situ*.

At present, research exploring the applications of biofilters is aimed at characterizing the functional properties of the antibody as it resides within the boundaries of the cell wall. Future efforts will be aimed at enhancing the functionality of antibodies in plants to enable catalytic processing of molecules retained within the cell. □

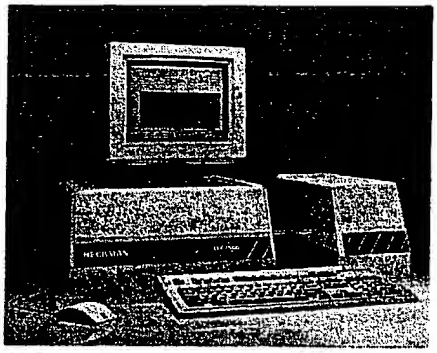
Andrew Hiatt is a Molecular Biologist at the Research Institute of Scripps Clinic, La Jolla, California 92037, USA. For more information, fill in reader service number 100.

1. Huse, W.D. *et al. Science* **243**, 1275–1281 (1989).
2. Hiatt, A.H., Cafferkey, R. & Bowdish, K. *Nature* **342**, 76–78 (1989).
3. Tramontano, A., Janda, K.D. & Lerner, R.A. *Science* **234**, 1566–1570 (1987).
4. Bevan, K.A., Binns, A.N., Matzke, A.J.M. & Chilton, M.D. *Cell* **32**, 1033–1043 (1983).
5. Christianson, M.L. & Warnick, D.A. *Dev. Biol.* **101**, 382–390 (1984).
6. Baumann, G., Raschke, E., Bevan, M. & Schoffl, F. *EMBO J.* **6**, 1161–1166 (1987).
7. Sturm, A., Van Kuik, A., Vliegenthart, J.F.G. & Chrispeels, M. *J. biol. Chem.* **262**, 13392–13403 (1987).
8. Varner, J.E. & Tuan-Hua Ho, D. in *Plant Biochemistry*, 3rd edn, 713–770 (Academic, New York, 1976).
9. Deom, C.M., Oliver, M.J. & Beachy, R.N. *Science* **237**, 389–394 (1987).
10. Yoder, O.C. *Ann. Rev. Phytopath.* **18**, 103–129 (1980).
11. Carlson, J.R. *Molec. cell. Biol.* **8**, 2638–2646 (1988).
12. Biocca, S., Neuberger, M.S. & Cattaneo, A. *EMBO J.* **9**, 101–108 (1990).
13. Bird, R.E. *et al. Science* **24**, 423–426 (1988).
14. Carpita, N., Sabularse, D., Montezinos, D. & Plemer, D. *Science* **205**, 1144–1147 (1979).

FASEB highlights

The Federation of American Societies for Experimental Biology (FASEB) annual meeting will be held in Washington, DC, next week. A micro-osmotic pump for slow-release drug delivery and a vertical tube gel apparatus will be among the many exhibits.

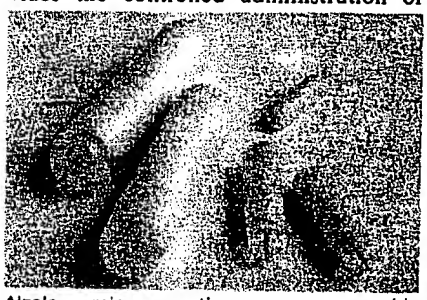
AT FASEB, Beckman Instruments will be launching the programmable DU7500 **diode array UV/visible spectrophotometer** designed for microvolume and ultramicrovolume samples of up to 100 µl (*Reader Service No. 101*). The patented



DU7500: Beckman's next generation of UV/visible spectrophotometers.

full spectrum quantitation applies all the data points in a scan to arrive at accurate and reliable component concentrations, says Beckman. Data are calculated using advanced vector quant maths. RediRead and RediScan modes allow the user to take readings or wavelength scans even when other measurements are in progress. A one-button prompt automatically sets up the new readings or scan, after which the interrupted research can be resumed. The DU7500 simplifies protein analysis by providing pre-selected parameters for Bradford (595 nm), Lowry (high sensitivity: 750 nm; low sensitivity: 500 nm), Biuret (540 nm) and direct UV method (280 nm). Kinetic analyses are run at single or multiple wavelengths: results can be displayed in five plot formats. Prices for the DU7500, which will be in action in booth 1312, range from $15,000–25,000 (US), depending on configuration and choice of accessories.

To meet demands for an **implantable mico-osmotic pump** that can deliver a variety of bioactive compounds to animals weighing less than 10 grams, Alza Corporation has introduced the Alzet Model 1007D (*Reader Service No. 102*). Measuring just 17 mm in length and weighing 350 mg when empty, the Model 1007D provides the controlled administration of



Alza's micro-osmotic pumps provide sustained-release drug delivery.

Production of Protein Pharmaceuticals in Transgenic Plants

Copy to mailroom 2/7/29

Philip C. Lyons
Plant Biotechnology Program, Texas A & M University
Institute of Biosciences and Technology, Houston, TX 77035, USA

Gregory D. May, Hugh S. Mason and **Charles J. Arntzen**
Plants and Human Health Program, Boyce Thompson Institute for Plant Research
Tower Road, Ithaca, NY 14853-1801, USA

Introduction

Genetic transformation techniques are now applied routinely to a large number of plant species. Through this technology, numerous genes that confer agronomically important traits such as pest resistance and herbicide tolerance have been introduced into plants. The successes of modern plant biotechnology exemplified by these achievements in crop improvement have created considerable interest in further exploiting the remarkable biosynthetic capacity of plants by developing transgenic plants that will produce valuable new products. With a number of recent demonstrations that transgenic plants are capable of producing functional foreign proteins and peptides of known or potential pharmaceutical importance (Table 1), new and exciting possibilities have been created in the ancient science of plant medicinal chemistry.

Biologically active peptides and proteins have many potential pharmaceutical applications, including use as vaccines, immunomodulators, growth factors, hormones, blood proteins and enzymes. Efforts to produce these compounds economically and in adequate quantities have become increasingly reliant upon the use of various prokaryotic and eukaryotic cell-culture expression systems. In this article, we will consider the use of transgenic plants as alternative eukaryotic expression systems for the production of recombinant protein pharmaceuticals, and the advantages of plants compared to other systems.

Expression in Transgenic Plants

For many reasons, transgenic plants are a feasible, and in some cases preferable, eukaryotic expression system for the production of valuable pharmaceuticals. The number of

Table 1

Peptides or proteins with known or potential pharmaceutical applications that have been expressed in transgenic plants.

Protein or Peptide	Application
Hepatitis B surface antigen	*Vaccine*
Norwalk virus capsid protein	*Vaccine*
Foot-and-mouth disease virus	*Vaccine*
Human rhinovirus 14	*Vaccine*
Human immunodeficiency virus	*Vaccine*
S. mutans surface protein	*Vaccine*
E. coli enterotoxin, B subunit	*Vaccine*
Malarial circumsporozoite epitopes	*Vaccine*
Mouse ZP3 protein epitope	*Vaccine*
Mouse catalytic antibody 6D4	*Antibody*
Mouse mAB Guy's 13	*Secretory Antibody*
Mouse Mab B 1-8	*Antibody*
Anti-phytochrome Fv protein	*Antibody*
Anti-substance P	*Antibody*
Human serum albumin	*Serum Protein*
Human protein C	*Serum Protein*
α-trichosanthin	*Cytotoxin*
Ricin	*Cytotoxin*
Human epidermal growth factor	*Neuropeptide*

plant species amenable to genetic transformation is now quite large [1,2]. These plants have the capacity to express foreign genes from a wide range of sources, including viruses, bacteria, fungi, insects, animals and other plants. Plants also are capable of high levels of protein expression; foreign protein concentrations as great as 30% of the total soluble protein have been reported in transgenic plants [3], and expression levels in excess of 1% are often attainable [4-8]. With rapid advances being made in the manipulation of foreign protein expression through the development of novel vectors and improved understanding of protein folding, assembly and processing in plants, one can expect that more consistent and higher levels of expression for a wider range of functional proteins and peptides will be attained.

Biologically active peptides and proteins have many potential pharmaceutical applications.

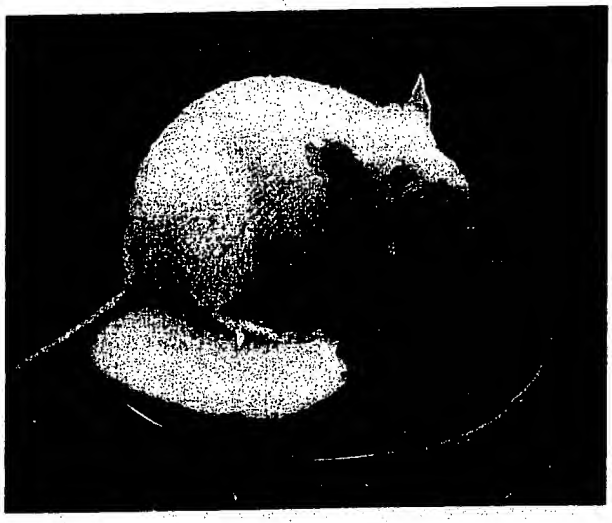
With respect to protein production, plants have an advantage of low cost involved in growing large amounts of biomass. With many of the difficulties traditionally encountered in achieving adequately purified protein from plants now overcome through advancements in protein purification techniques, it should now be possible to take greater advantage of plants as efficient sources of recombinant proteins. Furthermore, should it prove feasible to utilize transgenic plants as an edible source for the oral delivery of recombinant pharmaceuticals — as some evidence already suggests [9,10] — the need for costly purification procedures would in such instances be entirely eliminated.

Foreign proteins can be expressed in plants either transiently or as stably inherited traits [1,2,10]. Both approaches have been applied to the production of protein pharmaceuticals. Transient expression of pharmaceutically valuable proteins has been achieved most often by infection with a plant virus into which the gene of interest has been introduced through recombination [10]. Various modifications of this approach have been used with several different plant viruses. Infection with the genetically modified virus leads to the production of a fusion protein, typically consisting of the virus coat protein fused to a small foreign peptide sequence. Conversely, stable transformation is achieved by incorporation into the plant genome, by *Agrobacterium*-mediated transformation [1,2,10], of the foreign gene(s) encoding the protein of interest. Incorporation into the genome is the basis for the stability and heritability of traits introduced by this method.

Recombinant Vaccines in Transgenic Plants

Vaccines against viruses. A substantial number of candidate recombinant vaccines have been produced in plants through either stable or transient gene expression. Human virus diseases have been the most frequent target for plant-recombinant vaccines. The first of these was the hepatitis B surface antigen (HBsAg) [11]. Plant-derived recombinant HBsAg (rHBsAg) was expressed in tobacco plants (an important model plant species widely used for transgenic research). Although the expression levels were low (~0.01% of total soluble protein) it was shown in these experiments that the plant derived rHBsAg assembles into virus-like particles (VLPs). These particles were similar in size and other physical properties to yeast rHBsAg, which is the source of commercial hepatitis B vaccine (Recombivax®; Merck, Sharpe & Dohme). When a partially purified preparation of the tobacco-derived rHBsAg was parenterally injected into mice, it resulted in an antibody response that mimicked the response obtained with Recombivax [12]. In addition, lymph node T-cells isolated from mice, which were primed with tobacco-derived rHBsAg, could be stimulated to proliferate *in vitro* by both tobacco-derived and yeast-derived rHBsAg. These experiments demonstrate the close antigenic relatedness of the plant-derived recombinant protein to a known, highly effective recombinant vaccine produced in another eukaryotic expression system.

Another potential virus vaccine to be expressed in transgenic plants is the Norwalk virus capsid protein (NVCP) [13]. Norwalk virus causes epidemic acute gastroenteritis in humans. Plant-derived recombinant NVCP (rNVCP) has been expressed stably in both tobacco leaves and potato tubers, and like plant-derived rHBsAg, it self-assembled into VLPs. Expression and self-assembly of the capsid protein had first been achieved in recombinant baculovirus-infected insect cells [14] and this material has shown promise as an oral vaccine. In plant cells, the accumulation of rNVCP was nearly 0.3% of total protein; for potatoes, this resulted in a yield of about 20 µg rNVCP per gram of tuber weight. Feeding transgenic tubers expressing rNVCP to mice caused both humoral and mucosal anti-Norwalk virus antibodies to be produced. These results demonstrated that a plant-derived recombinant subunit antigen causes oral immunization when consumed as a food



Transgenic potatoes expressing a cholera vaccine have caused an oral immunization in mice fed the potatoes.

mutans, the principal cause of dental caries in humans. In this instance, four distinct components (a heavy and light chain, and two additional proteins, one of which is added during secretion) were expressed in a single transgenic plant. Furthermore, part of the heavy chain constant regions were replaced with domains from an IgA heavy chain since the secretory IgA form of antibody has proven more effective in passive immunotherapy against *S. mutans*. Remarkably, despite the heavy chain modification and the complexity of the recombinant molecule, all four proteins assembled into a functional antibody that recognized the native antigen from *S. mutans* and caused aggregation of cells of this bacterium. It was reported that efforts are now in progress to formulate the plant-derived recombinant antibody into toothpaste to determine if this is an effective means to prevent tooth decay.

A number of proteins present in human serum are of vital importance to medicine.

Other approaches have been used to produce antibodies in plants as well. In one instance [26], chimeric genes, containing the light or heavy chain of mAb B 1-8 fused with a plant signal peptide, were placed under the control of separate promoters in a single plasmid. Tobacco plants transformed with the chimeric genes produced both the light and heavy chains, which assembled into immunologically detectable antibodies that bound to the appropriate hapten. The level of expression was not determined in this study.

Another approach, first developed for antibody expression in *E. coli* [27], involves expression of single-chain Fv antibodies in which the light and heavy chain variable domains of an immunoglobulin are joined together by a flexible peptide linker. The flexible linker facilitates folding and assembly of the two chains, resulting in a functional synthetic antibody fragment. This approach was used successfully to express a functional anti-phytochrome single-chain Fv protein in transgenic tobacco [28]. It also was used with tobacco to express the antibody to a coat protein of the artichoke mottled crinkle virus [29], resulting in a reduced incidence of infection and delayed symptom development in virus-inoculated transgenic tobacco plants. Expression levels of the single-chain Fv antibodies in each instance was about 0.1% of total soluble leaf protein.

One additional approach used for antibody expression in plants involved the construction of a general purpose vector with a multiple cloning site that allows the insertion of a heavy chain variable (VH) domain [30]. VH domains expressed in *E. coli* have been shown to fold correctly and often retain antigen-binding activity. Using this vector, a "single-domain antibody" consisting of the VH domain of antibody to substance P (a neuropeptide) was expressed in transgenic tobacco, where it accumulated to approximately 1% of total soluble protein.

Serum Proteins in Transgenic Plants

A number of proteins (including coagulation and anti-coagulation factors) present in human serum are of vital importance to medicine. Due to the nature of donated plasma and the necessity for a sufficient supply of highly purified blood products, alternative production systems are required. Recombinant human serum albumin (rHSA) that is indistinguishable from the authentic human protein has been expressed in transgenic tobacco and potato [31]. The significance of this particular accomplishment rests not only in the demonstration that a valuable protein is produced in transgenic plants, but also that it was possible to achieve proper processing by fusion of rHSA to a plant presequence, resulting in cleavage and secretion of the correct protein. This achievement was particularly relevant in view of the substantial difficulties that have been encountered in other systems used to express recombinant HSA. The level of expression in transgenic potato plants was 0.02% of total soluble leaf protein.

Human serum protein C, a highly modified and glycosylated serine protease zymogen that requires proteolytic processing, also has been produced in transgenic tobacco plants (D.L. Weissenborn, Virginia Polytechnic Institute and State University, unpublished data). In this research, tobacco plants expressed both single chain and heavy chain forms of the protein, suggesting that the tobacco-derived recombinant protein C undergoes proper cleavage as well.

Recombinant Toxins in Transgenic Plants

Many toxins are proteins, usually derived from bacteria or plants, which kill cells by interfering with metabolism, often by the inhibition of protein synthesis. Due to the extreme cytotoxicity of many of these proteins, they have been the subject of intense investigation as tumor controlling agents. Most toxins exhibit very little site specificity, however, and in order to be used therapeutically, require modification of the binding domains so that the toxin will be preferentially directed to the appropriate target [32]. The modification usually involves inactivation or removal of the binding domain and expression of the toxin as a fusion protein with a vector that can deliver the toxin to the preferred site of action.

There have been two recombinant toxins expressed in transgenic plants to date [4,7], both of which are plant-derived eukaryotic ribosome-inactivating proteins (RIPs). The first of these, α-trichosanthin, a 27 kDa protein from *Trichosanthes kirilowii*, inhibits the replication of HIV in acutely infected $CD4^+$ lymphoid cells and in chronically infected macrophages. Although α-trichosanthin had been expressed previously in *E. coli*, the amount recovered was low, less than 0.01% of total cellular protein. Transient expression of this protein in tobacco plants using recombinant TMV as the vector resulted in accumulation of α-trichosanthin in leaves to at least 2% of total soluble protein 141.

Ricin is an RIP from castor bean that has therapeutic potential for the treatment of AIDS and cancer. Development of recombinant toxins utilizing ricin also has been hampered by low expression levels, as well as difficulties in achieving proper folding and processing in various expression systems.

However, stable expression of active, processed recombinant ricin was achieved in transgenic tobacco [7]. As with the transiently expressed α-trichosanthin, levels of the stably expressed recombinant ricin reached approximately 2% of the total soluble protein. The results obtained with these two toxins, together with the successful expression of fusion proteins described elsewhere in this article, suggest that transgenic plants may be quite useful for the production of site-selective toxins requiring fusion to a targeting protein.

Bioactive Peptides in Transgenic Plants

In addition to the vaccine epitopes discussed above, other small bioactive peptides with various potential pharmaceutical applications have been expressed in transgenic plants. Stable expressson of a chemically synthesized gene for human epidermal growth factor (hEGF), a small mitogenic peptide which stimulates *in vitro* proliferation of animal cells, was achieved in transgenic tobacco [33]. Although native hEGF is processed proteolytically in human cells from a larger precursor, the synthetic gene encoded only for the active peptide portion. Nevertheless, incorporation of this synthetic gene into the plant genome resulted in expression of a peptide that reacted with hEGF specific antibody. However, for plants to be feasible as a source of hEGF, expression levels will need to be increased, since the highest content of hEGF measured in transgenic plants was only 0.001% of total soluble proteins.

Leu-enkephalin, a pentapeptide from brain that exhibits opiate activity, has been produced using two different approaches. It was first expressed transiently in tobacco protoplasts as a fusion protein with the TMV coat protein, where it represented the major protein in the cells [34]. Subsequently, it was expressed as a fusion protein with a 2S albumin seed storage protein [35] in whole transgenic *Arabidopsis thaliana* and *Brassica napus* (oilseed rape) plants. Because 2S albumins represent up to 60% of total seed protein, depending upon the plant species, expression of foreign proteins in this manner should greatly facilitate protein purification. Using a 2S albumin-rich fraction from either species, it was possible to obtain highly purified peptide by tryptic digestion and HPLC. The yield obtained from *B. napus* was estimated to be equivalent to 15-75 g of peptide per hectare. Furthermore, the authors of this study suggested that by screening individual transformants for higher expression levels—by using lines of oilseed rape with higher protein/oil ratios and by standard breeding techniques—yields of the pentapeptide could be increased substantially.

An inhibitor of angiotensin I converting enzyme, found in the tryptic hydrolysate of milk, was expressed transiently in transgenic tobacco and tomato as a fusion protein with the TMV coat protein [9]. This 12-peptide inhibitor has anti-hypertensive effects when orally administered. Yield of the fusion protein in tomato fruit was approximately 10 μg fusion protein per g of plant tissue, equivalent to approximately 0.7 μg of peptide. It was suggested that the tomato fruit could be administered orally as a dietary antihypertensive agent, since the inhibitor should be released from the fusion protein in the intestine by trypsin digestion.

Conclusion

There is now ample evidence that transgenic plants are a feasible alternative system for the production of recombinant protein pharmaceuticals. From the examples that have been discussed here, it is apparent that this capability extends to a wide variety of pharmacologically active compounds. Although levels of accumulation of the recombinant proteins considered in this discussion varied considerably, from about 0.001% of total soluble protein to greater than 2%, in many instances, the levels achieved were comparable to, or greater than, that achieved with other systems. Furthermore, it is reasonable to expect that increased levels of production for protein pharmaceuticals are attainable, particularly since accumulation of non-pharmaceutical recombinant proteins to levels as great as 30% of total soluble protein has been reported. Particularly encouraging is the degree of success that has been attained in producing active, processed forms of the expressed candidate protein pharmaceuticals. Evidence for appropriate post-translational processing steps such as folding, assembly, secretion, and proper cleavage of precursor molecules indicate that even highly complex foreign proteins can be produced in plants and are likely to be functional.

There are certain aspects of transgenic plants beyond the usual considerations of protein expression levels that would, in some instances, make their use as an expression system preferable. For example, attempts to express RIPs and other cytotoxins to high levels in other systems have encountered difficulties because of their extreme toxicity. Plants, from which many cytotoxins are derived, are in some instances considerably less sensitive than other organisms. Another aspect concerns the prospect of delivering orally administered vaccines and other pharmaceuticals via edible plant tissues expressing these bioactive compounds. The eco-

Vegetables are known sources of nutraceuticals. The next step is to produce transgenic plants with novel pharmaceutical properties.

nomics and logistics of this approach are ideally suited to developing countries, where transportation and an adequate cold-chain — necessary for most current vaccines and many other drugs — are lacking.

The use of transgenic plants is, of course, not strictly limited to protein or peptide pharmaceuticals. In the future, more genes involved in the biosynthesis of non-protein pharmaceuticals will become available. Together with increased understanding of protein expression in plants, this should lead to greater reliance in the pharmaceutical industry upon transgenic plants for the production of both protein and non-protein drugs.

References

1. C.S. Gasser and R.T. Fraley, *Sci. Amer.*, 266, 62 (1992).
2. H. Uchimaya et al., *J. Biotechnology*, 12, 1 (1989).
3. K.E. McBride et al., *Proc. Natl. Acad. Sci. USA*, 91, 7301 (1994).
4. M.H. Kumagai et al., *Proc. Natl. Acad. Sci. USA*, 90, 427 (1993).
5. K.E. McBride et al., *BioTechnology*, 13, 362 (1995).
6. G.J.H. Van Rooijen and M.M. Moloney, *BioTechnology*, 13, 72 (1995).
7. P.C. Sehnke et al., *J. Biol. Chem.*, 269, 22473 (1994).
8. R.E. Shade et al., *BioTechnology*, 12, 793 (1994).
9. H. Hamamoto et al., *BioTechnology*, 11, 930 (1993).
10. H.S. Mason and C.J. Arntzen, *Trends Biotech.*, 13, 288 (1995).
11. H.S. Mason et al., *Proc. Natl. Acad. Sci. USA*, 89, 11745 (1992).
12. Y. Thanavala et al., *Proc. Natl. Acad. Sci. USA*, 92, 3358 (1995).
13. H.S. Mason et al., *Proc. Natl Acad. Sci. USA*, (in press).
14. X Jiang et al., *Science*, 250, 1580 (1990).
15. C.J. Arntzen et al., in *Vaccines '94, Modern Approaches to New Vaccines Including Prevention of AIDS*, 339 (1994).
16. R. Usha et al., *Virology*, 197, 366 (1993).
17. C. Porta et al., *Virology*, 202, 949 (1994).
18. L. McClain et al., *AIDS Res. Human Retrovirus*, 11, 327 (1995).
19. R. Curtiss and G.A. Cardineau, World Intellectual Property Organization PCT/US89/03799 (1990).
20 J.D. Clemens et al., *J. Infect. Dis.*, 158, 372 (1988).
21. T.A. Haq et al., *Science*, 268, 714 (1995).
22. T.H. Turpen et al., *BioTechnology*, 13, 53 (1995).
23. J. Fitchen et al., *Vaccine*, 13, 1051 (1995).
24. A. Hiatt et al., *Nature*, 342, 76 (1989).
25. J.K-C. Ma et al., *Science*, 268, 716 (1995).
26. K. During et al., *Plant Mol. Biol.*, 1S, 281 (1990).
27. R.E. Bird et al., *Science*, 242, 423 (1988).
28. M Owen et al., *BioTechnology*, 10, 790 (1992).
29. P. Tavladoraki et al., *Nature*, 366, 469 (1993).
30. E. Benvenuto et al., *Plant Mol. Biol.*, 17, 865 (1991).
31. P.C. Sijmons et al., *BioTechnology*, 8, 217 (1990).
32. R.J. Kreitman and I. Pastan, *Advances in Pharmacology*, 28, 193 (1994).
33. K. Higo et al., *Biosci. Biotech. Biochem.*, 9, 1477 (1993).
34. N. Takamatsu et al., *FEBS Lett.*, 269, 73 (1990).
35 J. Vandekerckhove et al., *BioTechnology*, 7, 929 (1989).

Charles J. Arntzen *is the President and CEO of the Boyce Thompson Institute for Plant Research, Inc., a not-for-profit corporation affiliated with Cornell University which is devoted to plant science. He was fomerly Dean and Deputy Chancellor for Agriculture at Texas A&M University.*

Hugh Mason*'s research focus in the past three years has been the use of plants as a recombinant expression system for foreign proteins, especially candidate vaccine antigens, and the testing of transgenic plant material for immunogenic activity in animals.*

Greg May *joined the Plants and Human Health Program at the Boyce Thompson Institute for Plant Research in December 1995. His current research efforts are directed at understanding the fundamental regulatory controls of a number of complex developmental plant processes.*

Phil Lyons *was at the Alkek Institute of Biosciences and Technology at the Texas Medical Center, where he was involved in a research project on the expression of antigenic proteins in transgenic plants.*

Comments? Questions? Critiques?

Send all correspondence to:
Pharmaceutical News
c/o 820 Town Center Drive
Langhorne, PA 19047, USA

Interested in contributing a technical review?

Contact:
John Vassallo
Pharmaceutical News
c/o 820 Town Center Drive
Langhorne, PA 19047, USA
Telephone: 215-750-2642; Fax: 215-750-6343

Transgenic Animals as Bioproducers of Therapeutic Proteins

Juhani Jänne[1,3], Juha-Matti Hyttinen[1,3], Teija Peura[2,3], Minna Tolvanen[2,3], Leena Alhonen[1,3] and Maria Halmekytö[1,3]

Many human therapeutic proteins are currently produced with the aid of recombinant DNA technology in microbial bioreactors and a few also in large-scale animal cell cultures. Although extremely cost-efficient, the microbial production system has many inherent limitations. Micro-organisms, such as bacteria, can read the universal genetic code and hence produce human proteins with correct amino acid sequence, but cannot carry out post-translational modifications, such as glycosylation, or fold the newly synthesized protein properly to ultimately generate a biologically active entity. Moreover, even though the production of the proteins as such is inexpensive, the downstream processing of the final product may be extremely difficult and costly. Many of these disadvantages, especially the lack of post-translational modifications, can be overcome by employing large-scale animal cell cultures for the production of proteins of pharmaceutical interest. However, due to the long generation time and the requirement for rich culture media, the use of animal cell bioreactors is unacceptably expensive. With the advent of transgenic technology, the production of human pharmaceuticals in large transgenic animals has become more and more attractive. The use of targeted gene transfer, the expression of the transgene of interest can be directed to occur in the mammary gland of large farm animals, such as pigs, sheep, goats or dairy cattle, and hence the transgene product is ultimately being secreted into the milk. Although not yet in commercial use, the last few years have witnessed a remarkable progress in this area and proved the feasibility of the use of 'molecular farming' in high-quantity, low-cost production of valuable therapeutic or industrial proteins. While reviewing the progress of the field over the past few years, we discuss in somewhat greater detail aspects connected with the use of dairy cattle as bioproducers of human therapeutic proteins.

Key words: transgenic animals; mouse; sheep; goat; cattle; bioreactor; pharmaceuticals; human erythropoietin; mammary gland; milk.

(Annals of Medicine 24: 273–280, 1992)

Transgenic animals are animals developed from embryos into which foreign genes have been transferred. If the foreign gene is introduced into the one-cell embryo (fertilized oocyte), and if integrated, the transgene becomes a dominant Mendelian genetic characteristic that is inherited by the progeny of the founder animal. The ability to genetically manipulate mammals has opened an immense potential with almost unlimited applications in basic and applied research.

From the [1]Department of Biochemistry & Biotechnology, [2]Department of Anatomy, and [3]A.I. Virtanen Institute, University of Kuopio, Finland.

Address and reprint requests: Dr Juhani Jänne, Department of Biochemistry & Biotechnology, University of Kuopio, P.O. Box 1627, SF-70211 Kuopio, Finland.

The experimental work of the author's laboratory has received financial support from the Ministry of Agriculture and Forestry and from the Academy of Finland.

Production of Transgenic Animals

There are many routes into the germ-line cells (for general reviews see refs 1–3). By far the most widely used is the microinjection of foreign genes into one of the two pronuclei of a fertilized oocyte. Typically, tens to hundreds of gene copies are injected into the pronucleus, of which some will become integrated (probably while repairing chromosomal damage). As a result of integration at the one-cell embryo stage, the foreign

gene potentially occurs in every cell of the born animal. The gene transfer can also be accomplished by retroviral infection of early embryos or transferring the transgene into embryonal stem cells followed by the introduction of the stem cells into blastocysts. All these methods are tedious and require sophisticated equipment (microinjection) or demanding cell culture techniques (embryonal stem cells). Therefore, it is no wonder that profound public interest was aroused by a report in 1989 (4) that foreign genes can be introduced into unfertilized oocytes by sperm cells in connection with *in vitro* fertilization. This observation was, with good reason, named as 'biological cold fusion' as the method would dramatically simplify the production of transgenic animals in general and large animals in particular. Unfortunately, in spite of serious attempts in a number of laboratories all over the world, this finding has so far not been reproduced (5) and hence the comparison with cold fusion appears to be correct.

The first transgenic mice produced by the microinjection technique were generated in 1980 (6) and since then hundreds of transgenic mouse lines have come into existence (1–3). Transgenic mice have established an exciting new experimental approach to mammalian molecular biology. Germ-line transmission of foreign genes makes it possible to study the function and regulation of genes during animal development in their 'natural' environment. More importantly, with the aid of transgenesis it is possible to construct animal models for almost every imaginable human disease ranging from tumourigenesis to a number of metabolic disorders (3). By using fusion gene constructs, a tissue-specific expression directed by various promoters can readily be achieved. Thus, transgenic mice also serve as important models for targeted gene transfer in large domestic animals.

The creation of 'giant' transgenic mice in 1982 (7) harbouring the metallothionein–rat growth hormone fusion gene in their genome not only attracted much public attention, but also led to the realization that transgenic techniques could be applied to the genetic engineering of domestic animals.

Transgenic Livestock

Even though various groups have produced transgenic farm animals, the progress in this field has not been nearly as dramatic as in the generation of transgenic mice. The reasons are obvious: the pregnancy is long, the litter size is small, the availability of fertilized eggs is limited, and technical difficulties are encountered in the microinjection technique. The latter include difficulties in visualizing the pronuclei due to deposition of opaque material in the cytoplasm of oocytes in some species (pigs and cattle). Nevertheless, a number of transgenic farm animals have been generated during the past few years. The giant mouse (7) apparently served as a model for the first transgenic livestock. With the reasonable assumption that insertion of extra copies of growth hormone genes would lead to accelerated growth, the first transgenic pigs harboured either bovine or human growth hormone genes driven by mouse metallothionein (liver-specific) promoter (8). The genes of human growth hormone-releasing factor and human insulin-like growth factor 1 were likewise used (8). Some of these animals indeed grew faster, converted feed more efficiently to body weight, and even had reduced backfat thickness (8). These desirable effects, however, were clouded by a long list of adverse effects: the animals suffered from lameness, lethargy and gastric ulcers; their glucose metabolism was severely disturbed with a striking elevation of plasma insulin (8); the reproductive capacity of pigs expressing the growth hormone transgenes was seriously impaired. As the health problems encountered in these transgenic pigs were believed to be caused by continuous exposure to high circulating levels of growth hormone, more specific gene constructs were used to generate transgenic pigs. A regulatable (by nutritional factors) fusion gene was constructed by fusing the promoter region of rat phosphoenolpyruvate carboxykinase to genomic bovine growth hormone structure gene (9). The resulting transgenic animals showed enhanced feed efficiency and decreased backfat thickness; however, negative characteristics included stress susceptibility, joint pathology and respiratory distress (9).

Transgenic sheep carrying metallothionein–growth hormone gene constructs have likewise been generated (10, 11). Marginal or no growth advantage in these animals was accompanied by serious health problems, such as diabetes and premature death (11).

It thus appears that by using various growth hormone gene constructs, any improvement of the quality of livestock can only be achieved at the cost of severe adverse effects. This is obviously related to the fact that growth hormone exerts an array of metabolic effects unrelated to its growth-promoting activity. It is highly likely that the metabolic effect cannot be entirely avoided even using strictly controllable promoters. Instead of modifying the normal physiology of the whole animal, a targeted, tissue-specific expression of the transgene seems a much more attractive approach. This applies especially to the mammary gland and to the genetic modification of the milk composition in large farm animals.

Genetic Modification of Milk Composition in Large Domestic Animals

Genetic modification of milk composition is based on the use of transgene constructs in which the structural gene of interest is driven by mammary gland-specific regulatory sequences. The milk protein genes appear to contain highly conserved regions, 'milk boxes', in their 5′ flanking regions (12–14). The milk protein genes are expressed exclusively in the mammary gland across the species boundaries. This is exemplified by a recent report (15) showing that the whey acidic protein gene, which only occurs in rodents and rabbits but not in pigs, is effectively expressed in the mammary gland of transgenic pigs. Thus, the expression of any structural transgene under the control of milk protein promoter can be

directed to the mammary gland in all likelihood without affecting the health of the animal.

There are two major approaches to genetic modification of milk composition. It is possible to improve milk quality by introducing more gene copies of milk proteins, to enhance bacterial resistance by transfer of the lysozyme gene, to break down lactose with the aid of lactase gene, or even to change the fat composition by introducing enzymes capable of changing saturated fats to unsaturated (reviewed in refs 14 and 16). The second approach is to create transgenic farm animals secreting proteins of pharmaceutical or industrial value into their milk. The feasibility of the latter approach, i.e. the expression of non-milk proteins under the control of regulatory sequences of a milk protein in the mammary gland, has been demonstrated by a large number of transgenic mouse models. Human tissue plasminogen activator (17), human urokinase (18), a human growth hormone (19) have been produced in the milk of transgenic mice, to mention a few examples. Similarly, human interleukin-2 has been produced in the milk of transgenic rabbits (20). Although the expression level of the human interleukin-2 gene (under the control of rabbit β-casein promoter) was relatively low, the authors suggested that lactating transgenic rabbits may be used as a source of human recombinant proteins, as the transgenic rabbits can be generated in a short time (20). In addition, the rabbit milk (100 g per female per day) has almost three times more protein content than cow's milk (20).

The approach of producing human therapeutic proteins was subsequently extended to include larger farm animals. Several transgenic sheep lines were generated in which human α1-antitrypsin (21) and antihaemophilic factor IX (22) were produced from fusion genes containing sheep β-lactoglobulin gene. Human factor IX is a good example of a human protein that is extensively modified (glycosylation and γ-carboxylation) post-translationally and therefore cannot be produced by bacterial fermentation. The transgenic sheep were produced by using ovine β-lactoglobulin gene with human α1-antitrypsin cDNA or human factor IX cDNA inserted into the 5′ untranslated region of the gene (21, 22). Both cDNAs were expressed, however, very inefficiently. The concentration of factor IX in the milk was 100,000 times lower than the level of endogenous β-lactoglobulin in sheep milk (22). Although α1-antitrypsin was expressed more efficiently (21), the concentrations were apparently much too low to be worth commercial exploitation. The low expression rate of the gene constructs were obviously attributable to the lack of introns, as the introns apparently greatly enhance transcriptional efficiency, at least in transgenic mice (23). As shown later, genomic sequences are expressed strikingly better in transgenic sheep.

In September 1991 three reports were published simultaneously, representing a real breakthrough in the field and demonstrating the feasibility of using farm animals as bioproducers of pharmaceuticals. The Edinburgh group (24) succeeded in generating transgenic sheep producing human α1-antitrypsin in their milk. Unlike their earlier efforts (21, 22), they now used genomic sequences of human α1-antitrypsin operationally linked to the ovine β-lactoglobulin promoter (24). This gene construct worked very well in transgenic mice generating up to 7 g/l of biologically active human α1-antitrypsin in the milk of the mice (25). Two of the transgenic founder sheep secreted human α1-antitrypsin into their milk at the level of 1–5 g/l and the third founder up to 35 g/l (24). In fact, human α1-antitrypsin was the major protein (nearly 50% of total protein) of the milk of the latter animal (24). It is noteworthy that the total milk protein content was almost twice that of normal sheep. The secretion of the recombinant protein remained at this high level throughout the lactation period with no signs of sustained lactation (24). The protein displayed full biological activity and was glycosylated like its plasma-derived counterpart (24). The authors also set up a conventional purification procedure based on ion exchange, dye affinity, hydrophobic interactions and molecular sieving, with the aid of which they achieved more than 95% purification of the protein (24). Finally, the transgenic animals were perfectly normal and healthy (24). This transgenic sheep line is currently in the process of commercialization.

Simultaneously with the report of the transgenic sheep, a U.S. group (26) published their production of transgenic goats harbouring the human tissue plasminogen activator gene (a mutated glycosylation variant) governed by murine whey acidic protein promoter. Two transgenic goats were born representing an integration rate of about 7% (26). The level of plasminogen activator in the milk was only 3 μg/ml, i.e. about 10% of that produced by a recombinant mouse cell line (26). However, the authors reported the birth of a further transgenic goat producing human tissue plasminogen activator under the control of β-casein promoter at levels that were three orders of magnitude higher than those in the first transgenic animals (26). The authors indicated that the level of human tissue plasminogen activator in the milk of the transgenic dairy goat could make it an economically viable bioreactor (26).

The same authors also exploited the feasibility of isolation and purification of the recombinant protein (plasminogen activator) from goat's milk (27). Using conventional purification procedures including acid fractionation, hydrophobic and immunoaffinity chromatography, they obtained, after 8000-fold purification, an apparently homogenous protein displaying 84% of the biological activity of mouse cell-derived plasminogen activator (27). It was calculated that the transgenic animal with the highest level of expression (3 mg/ml) would produce tissue plasminogen activator in 1 day's milk in quantities that are equivalent to a daily harvest of a 1000 l cell culture bioreactor (27). Nevertheless, the authors described the situation as 'a worse case scenario' as regards the expression level and the purification procedure (27).

The third report described the generation of transgenic dairy cattle harbouring the human lactoferrin cDNA in their genome (28). This transgene construct was driven by a 15 kbp-fragment of bovine αS1-casein (the most abundant protein in cow milk) 5′ flanking region and also containing some 3′ flanking region of the casein gene. The construct gave a transgenesis rate of 10% in

mice and about 20% in cattle (28). It is noteworthy that the cattle embryos were produced and fertilized *in vitro* using bovine ovaries obtained from slaughterhouses as the starting material (28). These three different transgenic farm animal species certainly prove the feasibility of this approach and some of them (the transgenic sheep) may even be commercially exploitable.

As we have been working in the production of transgenic dairy cattle, the rest of this article deals with the production of pharmaceuticals in the mammary gland of dairy cattle, with special reference to improvements that could make the whole procedure less labour-intensive and require a smaller number of recipient animals.

Production of Pharmaceuticals in the Bovine Mammary Gland

In Vitro *Maturation and Fertilization of Bovine Oocytes*

Like the Dutch group (28) we routinely collect cattle ovaries from the slaughterhouse, isolate the immature oocytes, mature them *in vitro* (with FSH, LH and oestradiol-17β) and carry out *in vitro* fertilization with frozen thawed bull semen (29, 30). The fertilized bovine oocytes, unlike mouse oocytes, are not transparent, therefore they have to be centrifuged briefly to visualize the pronuclei. The microinjections are carried out with a micromanipulator connected to a microscope equipped with differential interference contrast optics. Figure 1 shows the visualization of one of the pronuclei after displacement of the lipid material (dark) by centrifugation. After the microinjection, the embryos are cultured *in vitro* for a further 6–8 days before subsequent manipulations.

Gene Construct Used

The expression cassette we are currently using is outlined in Fig. 2. The construct is under the control of bovine αS1-casein promoter (about 2.2 kbp) containing the casein signal sequence fused with genomic sequences of the production gene (currently human erythropoietin) and followed by further 3′ flanking sequences of the αS1-casein. αS1-casein is the most abundant protein (10 g/l) (16) in cow milk and its promoter efficiently directs the expression of heterologous genes into the mammary gland of transgenic mice (18). We have generated five transgenic mouse lines harbouring the human erythropoietin gene under the control of bovine αS1-casein promoter. These mice express the construct in the mammary gland and to some extent in the salivary glands (unpublished results). We routinely generate the parts of the gene constructs by amplification of genomic DNA with the aid of polymerase chain reaction. This necessitates the sequencing of each new construct before use as the polymerase chain reaction and the subsequent cloning occasionally generate mutations that have to be corrected.

Sexing and Transgene Integration Analysis of Preimplantation Embryos

At late morula or early blastocyst stage the *in vitro* cultured embryos will be bisected with the aid of a micromanipulator. A biopsy of each embryo is subjected to sexing and transgene integration analysis. The sexing is based on the polymerase chain reaction method we recently adjusted specifically for the purpose of transgenic animals (30) and which so far has proved to be 100% accurate (31). The method to determine whether

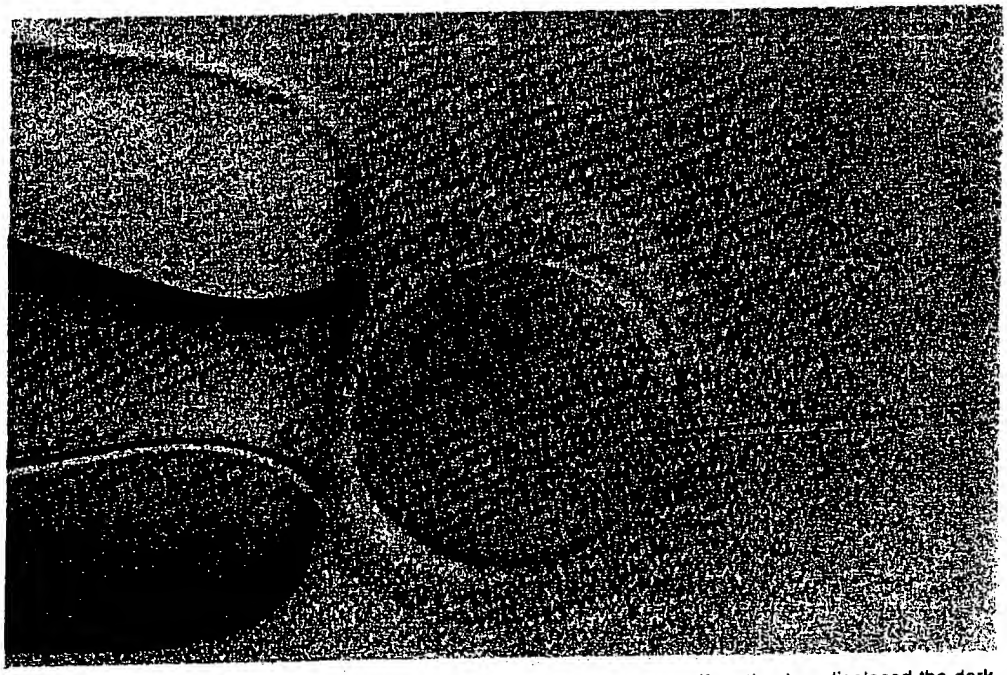


Figure 1. Microinjection of fertilized bovine oocyte. Note that a short centrifugation has displaced the dark cytoplasmic material, allowing visualization of one of the pronuclei.

the transgene has been integrated is based on the use of *Dpn* I restriction endonuclease. This is a unique enzyme as it only cleaves when an adenine belonging to its cleavage sequence (GATC) is methylated (32). The gene constructs are methylated before microinjection with bacterial *dam*-methylase (DNA adenine methylase), which methylates adenine residues at the sequence of GATC. As eukaryotic cells do not possess any maintenance methylase activity for adenine (eukaryotes methylate cytosine), the methylated adenine disappears after the integration and the subsequent replication of the transgene. When the demi-embryo is treated with *Dpn* I, non-integrated transgene is efficiently degraded (contains methyladenine in the cleavage sequence) but integrated and replicated transgene is not. The intact integrated transgene can be shown by using appropriate primers in the polymerase chain reaction. We are currently using a method in which the sexing and integration analysis are combined. A typical combined analysis is shown in Figure 3. Based on the integration analysis, our current integration rate for the human erythropoietin gene construct is about 20% for bovine embryos. After the sex and integration analyses, appropriate demi-embryos will be transferred non-surgically to synchronized recipient animals. The successful verification of the integration of the transgene before the embryo transfer will offer a new dimension for the generation of large transgenic farm animals as the need for recipient animals will be drastically reduced. This, combined with the almost unlimited availability of immature oocytes for maturation *in vitro*, makes the production of transgenic dairy cattle a realistic goal. The whole procedure, i.e. from slaughterhouse to recipient animals, is summarized in Figure 4.

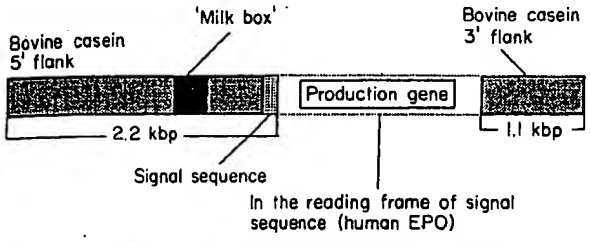


Figure 2. Fusion gene construct in which the structural gene (production gene) is under the control of bovine αS1-casein regulatory sequences. 'Milk box' refers to highly conserved sequences found in the 5' flanking region of milk protein genes in all mammals.

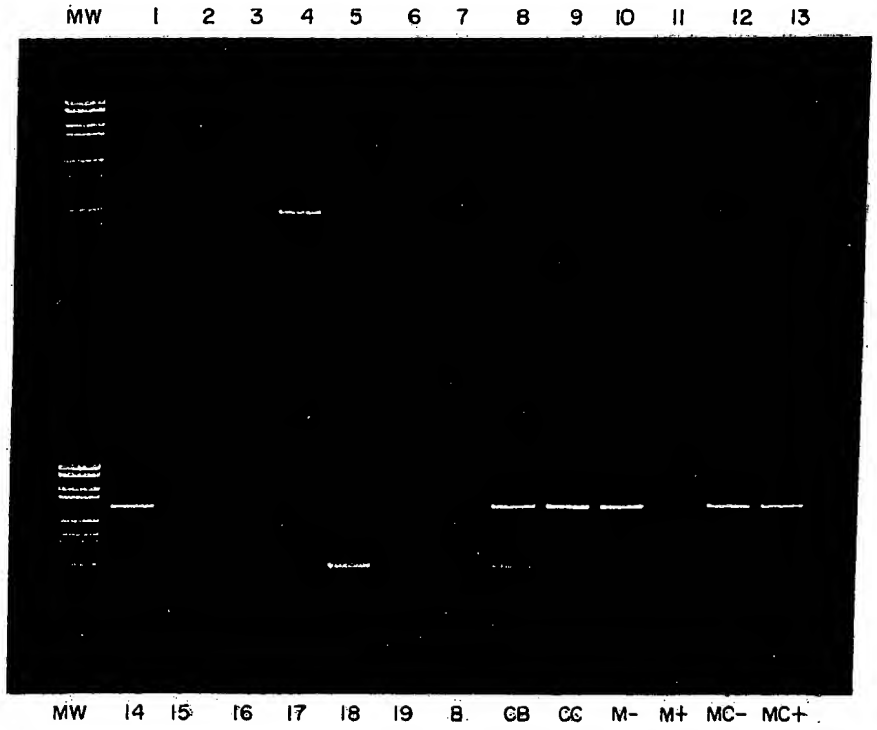


Figure 3. Transgene integration analysis and sex determination of bisected bovine embryos. The lanes are as follows: MW, molecular size markers; B, reagent blank; CB, DNA from transgenic mouse (human erythropoietin) mixed with DNA from bull. The upper fragment represents the transgene integration signal and the lower fragment the male-specific signal; CC, DNA from transgenic mouse (human erythropoietin) mixed with DNA from cow. Note that only the transgene signal is present; M−, microinjected human erythropoietin gene construct with *Dpn* I digestion; M+, microinjected gene construct with *Dpn* I digestion; MC−, DNA, from transgenic mouse mixed with microinjected gene construct without *Dpn* I digestion; MC+, as MC− but with *Dpn* I digestion. Lanes 1–19 represent analyses of bisected bovine embryos. Note that six embryos are positive for the transgene: two (male), three (female), five (male), 14 (female), 15 (male) and 16 (male).

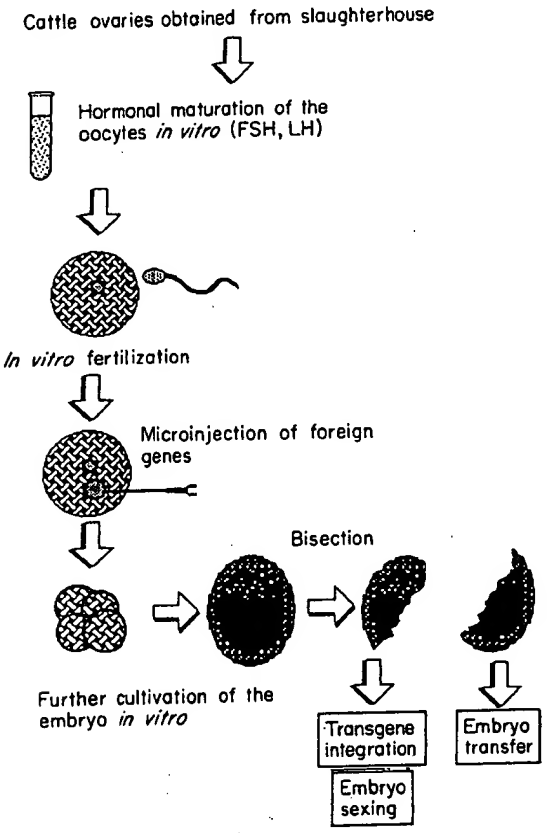


Figure 4. Production of transgenic cattle embryos from oocytes matured and fertilized *in vitro*. Before transfer into recipient animals, the embryos are bisected and one of the demi-embryos is used for transgene integration analysis and sex determination. Only embryos with established transgene integration and preferably of female sex will be transferred.

Possible Ways to Enhance and Control Transgene Expression

As already mentioned, and also experimentally demonstrated in mice (23), genomic sequences are much better transcribed than cDNAs. This probably also applies to large transgenic farm animals, as exemplified by the dramatic increase in the expression of human α1-antitrypsin in transgenic sheep following the replacement of cDNA (22) by genomic sequences (24) in the transgene construct.

The production of transgenic animals by microinjection technique leads to a random integration of the transgene in the embryonal genome and there is generally no direct relationship between the copy number of the integrated transgene and the level of expression (1). This lack of correlation is attributed to a chromosomal position effect that originates from a random integration of the transgene. The entirely unpredictable expression, if any, of a transgene may not be that important in case of transgenic mice, as new lines can be rapidly generated, but it is of utmost importance when large transgenic farm animals with a long pregnancy and small litter size are produced. Position-independent, gene copy number-dependent expression of transgene has been reported in a few cases, such as the human β-globin gene (33), the chicken lysozyme gene (34), the human CD2 gene (35), the human apolipoprotein E gene (36) and the human ornithine decarboxylase gene (37). The position-independent expression of the human β-globin and chicken lysozyme gene has been attributed to so-called matrix attachment elements (locus control regions or dominant control regions) flanking the 5′ and 3′ regions of the transgene and topologically sequestering a functional transcriptional unit (38, 39). It is generally believed that the locus control region is not an enhancer, but rather a new type of regulatory element influencing the organization of chromatin (40). The idea that a transgene construct could be 'shielded' by such regulatory elements has been recently tested using transgenic mice carrying the mammary gland-specific whey acidic protein gene. Interestingly, the inclusion of locus control region (origin not mentioned) in the gene construct led to an improved developmental regulation of the transgene and increased the proportion of the lines that expressed the transgene (41). The possibility of controlling, at least partially, the transgene expression understandably has a profound impact on the generation of large domestic transgenic animals.

Challenges and Opportunities

Although the production of transgenic bioreactors is still in its infancy and we eagerly await commercial applications, the approach certainly is viable. There are, however, many challenges and unsolved problems both concerning the generation of the transgenic bioproducers as well as the downstream processing of the milk to obtain the final product. Key issues in the production of the transgenic animal are the verification of transgene integration prior to the embryo transfer and the assurance of proper transgene integration assay using bisected preimplantation embryos. As indicated earlier, at least a partial solution for the latter problem may be the inclusion of matrix attachment elements in the transgene construct to protect the transgenic locus from a chromosomal position effect. In any event, the production of transgenic farm animals is a major effort both costwise and workwise, and hence it may be advisable to produce several transgene-derived products in the same animal. This can be accomplished by designing expression cassettes that contain two or more independent transcriptional units (structural genes driven by their own heterologous promoters). In fact, we are pursuing this direction by constructing an expression cassette containing transcriptional units for human erythropoietin and for one of the colony-stimulating factors.

As regards the downstream processing, an essential requirement for the transgene product is its stability in the milk. Certain evidence exists that an extensive proteolysis in the milk may not be a problem (24). Difficulties and low yields may also be encountered when purifying the transgene-derived product from the milk. Depending on the rate of expression, large purification factors and modest yields may be the result (27). The physiochemical properties of the milk, i.e. protein-fat

micelles, may likewise create extra problems in the design of proper purification procedures. The fact remains, however, that we are dealing with extremely high volumetric productions (especially dairy cattle) that permit lower purification yields.

The opportunities are almost unlimited. Using proper transgene constructs almost any imaginable protein of pharmaceutical or industrial interest can be produced in transgenic farm animals. Although the therapeutic proteins mentioned in this text are used for the treatment of relatively rare diseases, auxiliary indications for cytokines, for example, are emerging all the time. Thus, human erythropoietin or the colony-stimulating factors are not only used for the treatment of anaemia or leukopaenia but are also used as supportive treatment of cancer, AIDS, etc. (42). Moreover, the number of human therapeutic proteins currently available may look small, but new therapeutic peptides are being discovered continuously. Considering the yearly milk output of dairy cattle (6000–8000 l) and the milk content of αS1-casein (10 g/l), one cow carrying a transgene under the control of αS1-casein promoter would theoretically produce 60–80 kg/yr of the transgene-derived protein. In most cases, these proteins are therapeutically administered in microgram to milligram quantities. Despite having to compromise with modest yields and low expression rates, we would still be within the kilogram business. Add to this the advantages that the bioreactors feed and reproduce by themselves and that the transgenic lines have unlimited potential expansion as the transgene is inherited dominantly in a Mendelian fashion.

We thank Ms. Taru Koponen for her excellent secretarial assistance.

References

1. **Palmiter RD, Brinster RL.** Germ-line transformation of mice. *Ann Rev Genet* 1986; 20: 465–99.
2. **Jaenisch R.** Transgenic animals. *Science* 1988; 240: 1468–74.
3. **Gordon JW.** Transgenic animals. *Int Rev Cytol* 1989; 115: 171–229.
4. **Lavitrano M, Camaioni A, Fazio VM, Dolci S, Forage MG, Spadafori C.** Sperm cells as vectors for introduction of foreign DNA into eggs: genetic transformation of mice. *Cell* 1989; 57: 717–23.
5. **Brinster RL, Sandgren EP, Behringer RR, Palmiter RD.** No simple solution for making transgenic mice. *Cell* 1989; 59: 239–41.
6. **Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH.** Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* 1980; 77: 7380–4.
7. **Palmiter RD, Brinster RL, Hammer RE, Trubaner ME, Rosenfeld MG, Birnberg NC, Evans RM.** Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* 1982; 300: 611–5.
8. **Pursel VG, Hammer RE, Bolt DJ, Palmiter RD, Brinster RL.** Integration, expression and germ-line transmission of growth-related genes in pigs. *J Reprod Fert* 1990; Suppl 41: 77–87.
9. **Wieghart M, Hoover JL, McGrane MM, Hanson RW, Rottman FM, Holzman SH, Wager TE, Pinkert CA.** Production of transgenic pigs harbouring a rat phosphoenolpyruvate carboxykinase-bovine growth hormone fusion gene. *J Reprod Fert* 1990; Suppl 41: 89–96.
10. **Murray JD, Nancarrow CD, Marshall, JT, Hazelton IG, Ward KA.** Production of transgenic Merino sheep by microinjection of ovine metallothionein-ovine growth hormone fusion genes. *Reprod Fertil Dev* 1989; 1: 147–55.
11. **Rexroad CE, Hammer RE, Behringer RR, Palmiter RD, Brinster RL.** Insertion, expression and physiology of growth-regulating genes in ruminants. *J Reprod Fert* 1990; Suppl 41: 119–24.
12. **Hall L, Emery DC, Davies MS, Parker D, Craig RK.** Organization and sequence of the human a-lactalbumin gene. *Biochem J* 1987; 242: 735–42.
13. **Lubon H, Hennighausen L.** Nuclear proteins from lactating mammary gland bind to the promoter of a milk protein gene. *Nucl Acids Res* 1987; 15: 2103–21.
14. **Bremel RD, Yom H-C, Bleck GT.** Alteration of milk composition using molecular genetics. *J Dairy Sci* 1989; 72: 2826–33.
15. **Wall RJ, Pursel VG, Shamay A, McKnight RA, Pittius CW, Hennighausen L.** High-level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proc Natl Acad Sci USA* 1991; 88: 1696–700.
16. **Wilmut I, Archibald AL, Harris S, McClenaghan M, Simons JP, Whitelaw CBA, Clark AJ.** Modification of milk composition. *J Reprod Fert* 1990; Suppl 41: 135–46.
17. **Gordon K, Lee E, Vitale J, Smith AE, Westphal H, Hennighausen L.** Production of human tissue plasminogen activator in transgenic mice. *Biotechnology* 1987; 5: 1183–5.
18. **Meade H, Gates L, Lacy E, Lonberg N.** Bovine alpha-1-casein gene sequences direct high level expression of active human urokinase in mouse milk. *Biotechnology* 1990; 8: 443–6.
19. **Günzburg WH, Salmons B, Zimmermann B, Müller M, Erfle V, Brem G.** A mammary-specific promoter directs expression of growth hormone not only to the mammary gland, but also to Bergman glia cells in transgenic mice. *Mol Endocrinol* 1991; 5: 123–33.
20. **Bühler TA, Bruyère T, Went DF, Stranzinger G, Bürki K.** Rabbit β-casein promoter directs secretion of human interleukin-2 into the milk of transgenic rabbits. *Biotechnology* 1989; 8: 140–3.
21. **Simons JP, Wilmut I, Clark AJ, Archibald AL, Bishop JO, Lathe R.** Gene transfer into sheep. *Biotechnology* 1988; 6: 179–83.
22. **Clark AJ, Bessos H, Bishop JO, Brown P, Harris S, Lathe R, McClenaghan M, Prowse C, Simons JP, Whitelaw CBA, Wilmut I.** Expression of human anti-hemophilic factor IX in the milk of transgenic sheep. *Biotechnology* 1989; 7: 487–92.
23. **Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD.** Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci USA* 1988; 85: 836–40.
24. **Wright G, Carver A, Cottom D, Reeves D, Scott A, Simons P, Wilmut I, Garner I, Colman A.** High level of expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Biotechnology* 1991; 9: 830–4.
25. **Archibald AL, McClenaghan M, Hornsey V, Simons JP, Clark AJ.** High-level expression of biologically active human α1-antitrypsin in the milk of transgenic mice. *Proc Natl Acad Sci USA* 1990; 87: 5178–82.
26. **Ebert KM, Selgrath JP, DiTullio P, Denman J, Smith TE, Memon MA, Schindler JE, Monastersky GM, Vitale JA, Gordon K.** Transgenic production of a variant of human tissue-type plasminogen activator in goat milk: generation of transgenic goats and analysis of expression. *Biotechnology* 1991; 9: 835–8.
27. **Denman J, Hayes M, O'Day C, Edmunds T, Barlett C, Hirani S, Ebert KM, Gordon K, McPherson JM.** Transgenic expression of a variant of human tissue-type plasminogen activator in goat milk: purification and

characterization of the recombinant enzyme. *Biotechnology* 1991; 9: 839–43.
28. **Krimpenfort P, Rademakers A, Eyestone W, van der Schans A, van den Broek S, Koolman P, Kootwijk E, Platenburg G, Pieper F, Strijker R, de Boer H.** Generation of transgenic dairy cattle using '*in vitro*' embryo production. *Biotechnology* 1991; 9: 844-7.
29. **Peura T, Aalto J, Rainio V, Vanha-Perttula T.** Pregnancies from bovine oocyte matured and fertilized *in vitro*. *Acta Vet Scand* 1989; 30: 483–5.
30. **Peura T, Hyttinen J-M, Turunen M, Jänne J.** A reliable sex determination for bovine preimplantation embryos using the polymerase chain reaction. *Theriogenology* 1991; 35: 547–55.
31. **Peura T, Hyttinen J-M, Turunen M, Aalto J, Rainio V, Jänne J.** Birth of calves developed from embryos of predetermined sex. *Acta Vet Scand* 1991; 32: 283–6.
32. **Lacks S, Breenberg B.** A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. *J Biol Chem* 1975; 250: 4060–6.
33. **Grosveld F, van Assendelft GB, Greaves DR, Kollias G.** Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell* 1987; 51: 975–85.
34. **Bonifer C, Vidal M, Grosveld F, Sippel AE.** Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J* 1990; 9: 2843–8.
35. **Greaves DR, Wilson FD, Lang G, Kioussis D.** Human CD2 3′ flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell* 1989; 56: 979–86.
36. **Simonet WS, Bucay N, Lauer SJ, Wirak DO, Stevens ME, Weisgraber KH, Pitas RE, Taylor JM.** In the absence of a downstream element, the apolipoprotein E gene is expressed at high levels in kidneys of transgenic mice. *J. Biol Chem* 1990; 265: 10809–12.
37. **Halmekytö M, Alhonen L, Wahlfors J, Sinervirta R, Jänne OA, Jänne J.** Position-independent, aberrant expression of the human ornithine decarboxylase gene in transgenic mice. *Biochem Biophys Res Commun* 1991; 180: 262–7.
38. **Phi-Van L. Strätling WH.** The matrix attachment regions of the chicken lysozyme gene co-map with boundaries of the chromatin domain. *EMBO J* 1988; 7: 655–64.
39. **Stief A, Winer DM, Strätling WH, Sippel AE.** A nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature* 1989; 341: 343–5.
40. **Li Q, Zhou B, Powers P, Envert T, Stamatoyannopoulos G.** Primary structure of the goat β-globin locus control region. *Genomics* 1991; 9: 488–99.
41. **Wall RJ, McKnight RA, Shamay A, Hennighausen L.** Matrix attachment sequences improve genetic control of a mammary gland transgene in mice. *Theriogenology* 1992; 37: 319 (abstr).
42. **Niskanen E.** Hematopoetic growth factors in clinical hematology. *Ann Med* 1991; 23: 615–24.

Our Reference No. 9369-98

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: **Maurice Moloney, Joenel Alcantara and, Gijs van Rooijen**

Serial No. **09/402,488**

Filed: **February 16, 2000**

For: **Method For Cleavage of Fusion Proteins**

Group No.: 1656

Examiner: David J. Steadman

DECLARATION UNDER 37 C.F.R. 1.132

Honourable Assistant Commissioner
For Patents
Washington, D.C. 20231

Sir:

I, Maurice Moloney, citizen of Ireland and resident of Calgary, Alberta, Canada, hereby declare and state as follows:

1. I am one of the named inventors of the above-referenced patent application (hereinafter "the application").

2. I am the Chief Scientific Officer at SemBioSys Genetics Inc., the owner of the application. I have been involved in research relating to plant molecular biology for over 20 years. I attach a copy of my curriculum vitae as Exhibit A.

3. I have reviewed the Official Action issued on December 30, 2005, in connection with the application. In particular, I note the rejection of claims 1, 4, 6-9, 13, 15 and 19 under 35 USC §103(a), claim 5 under 35 USC §103(a), claims 10, 16 and 48-49 under 35 USC §103(a) and claims 14 and 50 under 35 USC §103(a).

4. Pages 10-15 of the Office Action discuss the teachings of Ward *et al.* (US Patent 6,265,204), Walsh *et al.*, (J. Biotech 45:235-241) and Yonezawa *et al.* (Int J Pept Protein Res 47:56-61). At page 11, the Office Action cites these references in support of the position that:

> one of ordinary skill in the art would have recognized that chymosin is an appropriate endoproteinase for cleaving a fusion protein at a Phe-Met junction of a desired protein with an N-terminal chymosin pro-peptide.

At pages 14-15, the Office Action notes (i) Walsh's teaching of "a bovine kappa-casein chymosin cleavage site (cleavage between Phe-Met, wherein Phe is at the P1 position and Met is at the P1' position) as a cleavable linker in a fusion protein," (ii) knowledge in the art that "pro-peptides of chymosin from mammalian and fungal sources also have a Phe at the P1 position," and (iii) knowledge that "most" heterologous proteins have "a Met at the N-terminus," and alleges that a skilled artisan would be motivated to use a chymosin pro-peptide as a fusion protein linker as claimed in the application, because such a fusion protein would have a Phe-Met junction that allegedly would be specifically cleaved by chymosin. Those assertions are not supported by the cited references, however, and contradict knowledge in the art regarding the activity of chymosin.

5. I have reviewed the Ward, Walsh and Yonezawa references cited in the Office Action. I have also reviewed Visser *et al.*, Biochim Biophys Acta 438: 265-72 (1976), which is cited by Walsh, and Schattenkerk *et al.*, Recl. Trav. Chim. Pays-Bas 90: 1320-22, (1970), which is cited by Visser. These references (attached as Exhibits B and C)

contradict the assertion that chymosin can be used to cleave a fusion protein at any Phe-Met junction.

6. These references do teach that chymosin cleaves the substrate κ-casein at a Phe-Met bond, but they also teach that the primary structure of the amino acids surrounding the Phe-Met bond is essential to the cleavage reaction. For example, Walsh *et al.* teaches at page 239, second column, that "[s]pecific hydrolysis of Phe_{105}-Met_{106} of κ-casein at pH 6.8 by chymosin is dependent upon the composition and sequence of amino acid residues in an extended region of the primary structure." (Emphasis added.) Consistent with Walsh, the abstract and page 271 of Visser *et al.* teach that "the sequence -Ser-Phe-Met-Ala- with a further residue added to either end is necessary to induce any cleavage by the enzyme." Figure 1 (page 1321) of Schattenkerk *et al.* provides further data on this point, showing that methyl esters (used for their solubility) of the following κ-casein-based peptides were completely resistant to cleavage by chymosin (referred to as "rennin" in the paper):

Phe-Met

Phe-Met-Ala

Phe-Met-Ala-Ile

Phe-Met-Ala-Ile-Pro-Pro-Lys-Lys

Ser-Phe-Met

Ser-Phe-Met-Ala

Leu-Ser-Phe-Met

Schattenkerk also reports that the peptide Leu-Phe-Met-Ala was completely resistant to cleavage by chymosin. On the other hand, Schattenkerk teaches that methyl esters of other κ-casein-based peptides (*e.g.*, Leu-Ser-Phe-Met-Ala-Ile) exhibited 99% cleavage by chymosin. As summarized by Visser *et al.* on page 271, second paragraph, "it is apparent that a minimum chain length of five amino acids residues including the sequence -Ser-Phe-Met-Ala- is essential to bring about any cleavage of the Phe-Met bond."

7. These teachings contradict the assertion in the Office Action that "chymosin is an appropriate endoproteinase for cleaving a fusion protein at a Phe-Met junction of a desired protein with an N-terminal chymosin pro-peptide," because they show that chymosin does not cleave substrate proteins at any and all Phe-Met junctions.

8. I also draw the attention of the Examiner to the Examples provided in the current application (USSN 09/402,288). As seen in Example 1, and specifically in Figure 1, the deduced amino acid sequence of the GST-Pro-Hirudin fusion protein includes a Phe-Met bond between amino acids 153 and 154 that is not cleaved during the cleavage reaction. This is consistent with the findings of Visser *et al.* and Walsh *et al.* which demonstrate the need for additional amino acids in order for the Phe-Met bond to be a substrate for chymosin. In addition, the cleavage site of the fusion protein in Figure 1 is not between a Phe-Met bond, but is between a Phe-Val bond between amino acid residues 278 and 279. Likewise, the cleavage site of the His-Pro-cGH fusion protein in Figure 2 is not between a Phe-Met bond, but is between a Phe-Ser bond between amino acid residues 84 and 85.

9. In summary, while the prior art teaches that chymosin cleaves the substrate κ-casein at a Phe-Met bond, the prior art and the examples in the application demonstrate that chymosin does not cleave substrate proteins at any and all Phe-Met bonds that are present. Thus, the basis for the obviousness rejections set forth in the Office Action is contradicted by the art of record and the additional prior art references discussed above.

10. At pages 10-11, the Office Action states that "Methionine is usually the first amino acid of a given polypeptide." This assertion is not consistent with the knowledge in the art regarding the primary structure of polypeptides. While the methionine codon is the "start" codon, a vast number of functional proteins do not have methionine at

their N-terminus because they are derived *in vivo* from fusions with signal sequences, transit peptides and other labile linkers and undergo modifications which eliminate the N-terminal methionine. Thus, a recombinant polypeptide of interest will not necessarily have an N-terminal methionine, and a prochymosin-polypeptide junction will not necessarily have a Phe-Met sequence.

11. I also point out that invention described and claimed in the application would not have been expected from the references discussed above, because they specifically teach that the sequence -Ser-Phe-Met-Ala- is essential for the reported chymosin cleavage, and do not indicate that chymosin can cleave a substrate protein at a Phe-Val or Phe-Ser bond, as shown in the examples of the application.

12. I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and, further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such a willful false statement may jeopardize the validity of the application or any patent issuing thereon.

April 7, 2006
Date

Maurice Moloney

Exhibit A

Dr. Maurice Moloney

SemBioSys Genetics Inc
110, 2985-23rd Avenue NE
Calgary, Alberta T1Y 7L3
(403) 250-5424 x758 Fax: (403) 250-3886
Email: moloneym@sembiosys.com

Dr. Maurice Moloney is the scientific founder of SemBioSys and serves as the company's chief scientific officer. He is also a professor in the Department of Biological Sciences at the University of Calgary and holds the Natural Sciences and Engineering Research Council of Canada (NSERC) Industrial Research Chair in plant biotechnology. Dr. Moloney's career in plant technology spans 20 years. Prior to founding SemBioSys, he spent seven years in his post at the University of Calgary, pursuing research on seed-specific gene expression, herbicide resistance and the plant cell cycle. Previously, he was the head of the Cell Biology Group at Calgene Inc., where he developed the first transgenic oilseed plants using Canola as the model. This resulted in a landmark patent in plant biotechnology and eventually became the basis of RoundUp Ready and Liberty Link Canola.

Dr. Moloney has published more than 70 original research papers and is an inventor on 22 issued or pending patent families. He serves on the advisory board of the National Research Council's Plant Biotechnology Institute and two other biotechnology companies. He serves on numerous Federal government committees including NSERC Council (the governing body of NSERC), CFI, National Research Council (Canada) and the Networks of Centres of Excellence programs. He was the co-president of the International Society for Plant Molecular Biology (ISPMB) Congress in 2000 and serves on the ISPMB board. Dr. Moloney has received a number of prestigious awards, including the Alberta Science and Technology (ASTECH) Award for leadership in Alberta Technology. He holds a Bachelor of Science degree in organic chemistry from Imperial College, University of London, and was awarded his doctorate in plant biochemistry from De Montfort University/Leicester Polytechnic in the United Kingdom.

Academic, Research and Industrial Appointments:

2004	Doctor of Science (*honoris causa*) University of Lethbridge
2003 – present	Member, Executive Committee of NSERC
2003 - present	Chair, Committee on Research Partnerships - NSERC
2002 - present	Appointed to the Natural Sciences and Engineering Research Council of Canada (NSERC) for a three year term
1999	ASTECH Award for leadership in Alberta Technology.
1995 – present	Professor, NSERC/Dow AgroSciences Industrial Research Chair of Plant Biotechnology, University of Calgary, Dept. of Biological Sciences
1994 – present	Founder and Chief Scientific Officer, SemBioSys Genetics Inc.
1990 – 1995	Associate Professor, University of Calgary, Dept. of Biological Sciences
1986 – 1990	Assistant Professor, University of Calgary, Dept. of Biological Sciences
1983 – 1986	Principal Scientist and Coordinator, Calgene Inc., Cell Biology Group
1979 – 1983	Royal Society European Postdoctoral Fellow, University of Lausanne, Institut de Physiologie et de Biologie Vegetales

1976 – 1979 Research Assistant, Leicester Polytechnic

Education:

1979 Doctorate in Plant Physiology, Leicester Polytechnic
1974 Bachelor of Science (Chemistry), Imperial College, University of London, UK

Research Contributions:

Articles in refereed publications (since 1995):

Abell, B.M., Hahn, M., Holbrook, L.A., **Moloney, M.M.** (2004) Membrane Topology and Sequence Requirements for Oil Body Targeting of Oleosin. The Plant Journal. 37: 461-70

Kühnel B, Alcantara, J., Boothe, J., van Rooijen, G., **Moloney, M.M.** (2003) Precise and Efficient Cleavage of Recombinant Fusion Proteins Using Mammalian Aspartic Proteases. Protein Engineering. Protein Engineering, Vol. 16 (10) pp. 777-783.

Seon, J.H., Szarka, S.J.., **Moloney, M.M.** (2002) A Unique Strategy for Recovering Recombinant Proteins from Molecular Farming: Affinity Capture on Engineered Oilbodies Journal of Plant Biotechnology 4(3) pp. 95 – 101.

Abell, B. M., High, S., and **Moloney, M.M.** (2002) Membrane Protein Topology of Oleosin is Constrained by its Long Hydrophobic Domain. Journal of Biological Chemistry, 277 (10) 8602-8610.

Nykiforuk, C.L., Furukawa-Stoffer, T.L., Huff, P.W., Sarna, M., Laroche A., **Moloney M.M.**, Weselake, R.J. (2002) Characterization of cDNAs Encoding Diacylglycerol Acyltransferase from Cultures of *Brassica napus* and Sucrose-Mediated Induction of Enzyme Biosynthesis. Biochimica et Biophysica Acta 1580: 95-109.

Moloney, M.M. (2000) Molecular Farming Using Seeds as Hosts. Seed Technology and its Biological Basis, Sheffield Academic Press, p. 226-256.

Weselake, R.J., Nykiforuk, C.L., Larochet, A., Patterson, N.A., Wiehler, W.B., Szarka, S.J., **Moloney, M.M.**, Tari, L.W., Derekh, U. (2000) Expression and Properties of Diacylglycerol Acyltransferase from Cell-suspension Cultures of Oilseed Rape. Biochemical Society Transactions 28(6): 684-686.

R.J. Weselake, A. Laroche & **M.M. Moloney** (2001) Molecular strategies for increasing oil content in canola. Final Report for Farming for the Future (AARI), project # 980810, 139 pages.

Moloney MM (2000) Seeds as Repositories of Recombinant Proteins in Molecular Farming. Korean Journal of Plant Tissue Culture 27 (4): 283-297.

Crowe A J, Abenes, M., Plant, A., **Moloney, M.M.** (2000) The seed-specific transactivator, ABI3, induces oleosin gene expression. Plant Science 151:171-181.

Deckers H, **Moloney MM** and Baum A (1999) The case for recombinant production of pharmaceutical proteins in plants. Annual Reports in Medicinal Chemistry 34: 237-245

Hays DB, Wilen RW, Sheng C, **Moloney MM** and Pharis RP (1999) Embryo-specific gene expression in microspore-derived embryos of *Brassica napus*. An interaction between abscisic acid and jasmonic acid. Plant Physiology 119: 1065-1072

Saborio F, Moloney MM, Tung P and Thorpe TA (1999) Root induction in *Pinus ayacahuite* by co-culture with *Agrobacterium tumefaciens* strains. Tree Physiology 19: 383-389

Ward KA, Holbrook LA, Lamb N, Abrams S, Reid DM and **Moloney MM** (1999) Structural requirements for biologically active jasmonates: Induction of protease inhibitors and seed specific proteins. Plant Growth Regulation 27: 49-56

Chaudhary S, Parmenter DL and **Moloney MM** (1998) Transgenic *Brassica carinata* as a vehicle for the production of recombinant proteins in seeds. Plant Cell Reports 17: 195-200

Kathiresan A, Nagarathna KC, **Moloney MM**, Reid DM and Chinnappa CC (1998) Differential regulation of 1-aminocyclopropane-1-carboxylate synthase gene family and its role in phenotypic plasticity in *Stellaria longipes*. Plant Molecular Biology 36: 265-274

Kermouni A, Mahmoud SS, Wang S, **Moloney MM** and Habibi HR (1998) Cloning a full-length insulin-like growth Factor-I complementary DNA in the goldfish liver and ovary and development of quantitative PCR method for its measurement. Gen Comp Endocrin 111: 51-60

Layzell DB, Brisson N, Devine MD, **Moloney MM**, Taylor GJ, Timmer V, Yada RY and Wood K (1998) Plant biology and food science in Canada: a vision for the future. Can J Bot. 76: 355-364

Mahmoud SS, Wang S, **Moloney MM** and Habibi HR (1998) Production of biologically active recombinant goldfish type II growth hormone in *Escherichia coli.* Comparative Biochemistry and Physiology 120: 657-663

Abell BM, Holbrook LA, Abenes M, Murphy DJ, Hills MJ and **Moloney MM** (1997) Role of the Proline Knot Motif in Oleosin Endoplasmic Reticulum Topology and Oil Body Targeting. The Plant Cell 9: 1481-1493

Abenes M, Holbrook LA, **Moloney MM** (1997) Transient expression and oil body targeting of an *Arabidopsis* oleosin-GUS reporter fusion protein in a range of oilseed embryos. Plant Cell Reports 17: 1-7

Holbrook LA, Tung P, Ward KA, Reid DM, Abrams S, Lamb N, Wilson Quail J, **Moloney MM** (1997) Importance of the chiral centres of jasmonic acid in the responses of plants: activities and antagonism between natural and synthetic analogues. Plant Physiology 114: 419-428

Liu J-H, Selinger LB, Hu Y-J, **Moloney MM**, Cheng K-J and Beauchemin KA (1997) An endoglucanase from the anaerobic fungus *Orpinomyces joyonii*: characterization of the gene and its product. Canadian Journal of Microbiology 43: 477-485

Liu J-H, Selinger LB, Cheng K-J, Beauchemin KA and **Moloney MM** (1997) Plant seed oil-bodies as an immobilization matrix for a recombinant xylanase from the rumen fungus *Neocallimastix patriciarum.* Molecular Breeding 3:463-470.

Markley NA, Young D, Laquel P, Castroviejo M and **Moloney MM** (1997) Molecular genetic and biochemical analysis of *B. napus* PCNA function. Plant Molecular Biology 34: 693-700

Hays DB, Rose P, Abrams SR and **Moloney MM** (1996) Biological activity of optically pure C-1 altered abscisic acid analogs in *Brassica napus* microspore embryos. J Plant Growth Regul 15: 5-11

Kühnel B, Holbrook LA, **Moloney MM** and van Rooijen GJH (1996) Oil bodies of transgenic *Brassica napus* as a source of immobilized ß-glucuronidase. JAOCS 73: 1533-1538

Mahmoud SS, **Moloney MM** and Habibi HR (1996) Cloning and sequencing of the goldfish growth hormone cDNA. Gen Comp Endocrin 101: 139-144

Zhang X-H, **Moloney MM** and Chinnappa CC (1996) Analysis of an ABA- and osmotic stress inducible dehydrin from *Stellaria longipes.* Journal of Plant Physiology 149: 617-622

Parmenter DL, Boothe JG, van Rooijen GJH, Yeung EC and **Moloney MM** (1995) Production of biologically active hirudin in plant seeds using oleosin partitioning. Plant Mol Biol 29: 1167-1180

Szarka SJ, Fitch M, Schaerer S and **Moloney MM** (1995) Classification and expression of a family of cyclin gene homologues in *Brassica napus.* Plant Mol Biol 27: 263-275

van Rooijen GJH and **Moloney MM** (1995) Plant seed oil-bodies as carriers for foreign proteins. BioTechnology 13: 72-77

van Rooijen GJH and **Moloney MM** (1995) Structural requirements of oleosin domains for subcellular targeting to the oil body. Plant Physiology 109: 1353-1361

Refereed Reviews or Book Chapters:

Moloney, MM (2002) Oleosin Partitioning Technology for Production of Recombinant Proteins in Oil Seeds **In** *Handbook of Industrial Cell Culture: Mammalian, Microbial, and Plant Cells* Eds. Vinci, Victor A., Parekh, Sarad R. Humana Press p. 279 – 298.

Moloney, MM (2002) Plant Molecular Farming: Using Oleosin Partitioning Technology in Oilseeds **In** *Plants as Factories for Protein Production* Eds. Hood, E., Howard, J. Kluwer Academic Publishers, Dordrecht, The Netherlands p. 55 – 75.

Moloney MM (2000) Molecular farming using seeds as hosts. *In* Seed Technology (Eds Bewley D and Black M) Sheffield Academic Press. pp 226-253

Moloney MM (1998) Oleosins as carriers for foreign protein in plant seeds. *In* Engineering crops for industrial end uses (Eds Shewry PR, Napier, JA, Davis P) Portland Press pp 47-54

Moloney MM and Holbrook LA (1997) Subcellular targeting and purification of recombinant proteins in plant production systems. *In* Biotechnology and Genetic Engineering Reviews (Ed MP Tombs) pp 14:321-336

Moloney MM, van Rooijen GJH (1996) Recombinant proteins via oleosin partitioning. INFORM 7(1) 107-114

Parmenter DL, Boothe JG **Moloney MM** (1996) Production and purification of recombinant hirudin from plant seeds. *In* Transgenic plants: a production system for industrial and pharmaceutical proteins (Eds Owen MRL, Pen J) John Wiley & Sons

Moloney MM (1995) "Molecular farming" in plants: achievements and prospects. *In* Biotechnol & Biotechnol Equipment 9/1995/1 3-9

Patents:

Moloney, MM, Boothe J, van Rooijen G. (2005) Oil Bodies and Associated Protein as Affinity Matrices. United States Patent No. 6,924,323, issued August 2, 2005.

Deckers, H, van Rooijen, G, Boothe, J, Goll, J, **Moloney, MM**, Schryvers, AB, Alcantara, J, Hutchins, WA (2004) Immunogenic Formulations Comprising Oil Bodies. United States Patent No. 6,761,914, issued July 13, 2004.

Moloney, MM, van Rooijen, G. (2004) Preparation of Heterologous Proteins on Oil Bodies. United States Patent No. 6,753,167, issued June 22, 2004.

Moloney, MM, Dalmia, BK (2004) Preparation of Thioredoxin and Thioredoxin Reductase Proteins on Oil Bodies. United States Patent No. 6,750,046, issued June 15, 2004.

Peoples, OP, **Moloney, MM**, Patterson, N, Snell, KD (2003) Modification of Fatty Acid Metabolism in Plants. Publication No. 2003/0233677, published December 18, 2003.

Moloney, MM, van Rooijen, G. Expression of Epidermal Growth Factor in Plant Seeds. Publication No. US 2003/0177537, published Sept. 18, 2003.

van Rooijen, G., Alcantera, J., **Moloney, MM** (2003) Method for cleavage of fusion proteins. Publication No. US 2003/0166162, published September 4, 2003.

Deckers, HM, van Rooijen, G, Boothe, J, Goll, J, **Moloney, MM** (2003) Products for Topical Applications Comprising Oil Bodies. United States Patent No. 6,599,513, issued July 29, 2003.

Deckers, HM, van Rooijen, G, Boothe, J, Goll, J, **Moloney, MM** (2003) Products for Topical Applications Comprising Oil Bodies. United States Patent No. 6,596,287, issued July 22, 2003.

Moloney, MM, van Rooijen, G, (2003) Preparation of Heterologous Proteins on Oil Bodies. United States Patent No. 6,753,167, issued June 22, 2004.

Peoples, OP, **Moloney, MM**, Patterson, N, Snell, KD (2003) Modification of Fatty Acid Metabolism in Plants. United States Patent No. 6,586,658, issued July 1, 2003.

Deckers, HM, van Rooijen, G, Boothe, J, Goll, J, **Moloney, MM** (2003) Products for Topical Applications Comprising Oil Bodies. United States Patent No. 6, 582, 710, issued June 24, 2003.

Moloney, MM, Boothe, J, van Rooijen, G. (2003) Oil Bodies and Associated Protein as Affinity Matrices. Publication No. US 2003/0096320, published May 22, 2003.

Szarka, S, **Moloney, MM** (2003) Method for production of multimeric immunoglobulins and related compositions. Publication No. US 2003/0093832, published May 15, 2003.

Moloney, MM, Boothe, J, van Rooijen, G. (2003) Oil Bodies and Associated Protein as Affinity Matrices. Publication No. US 2003/0059910, published Mar. 27, 2003.

Moloney, MM, Boothe,J., van Rooijen, G. (2003) Oil Bodies and Associated Protein as Affinity Matrices. United States Patent No. 6,509,453, issued January 21, 2003.

Deckers, HM, van Rooijen, G, Goll, J., **Moloney, MM** (2002) Products for topical applications comprising oil bodies. United States Patent No. 6,372,234, issued April 16, 2002.

Deckers HM, van Rooijen G, Boothe J, Goll J, Mahmoud S, **Moloney MM** (2001) Uses of Oil Bodies. United States Patent No. 6,210,742, issued April 3, 2001.

Habibi H, **Moloney MM** (2001) Expression of Somatotropin in Plant Seeds. United States Patent No. 6,288,304, issued September 11, 2001.

Deckers HM, van Rooijen G, Boothe J, Goll J, and **Moloney MM** (2001) Oil Body Based Personal Care Products. United States Patent No. 6,183,762 B1, issued Feb. 6, 2001.

Deckers HM, van Rooijen G, Boothe J, Goll J, Mahmoud S and **Moloney MM** (2000) Uses of Oil Bodies. United States Patent No. 6,146,645, issued Nov. 14, 2000.

Cheng K-J, Selinger LB, Liu J-H, Hu Y, Forsberg CW and **Moloney MM** (2000) Xylanase obtained from an anaerobic fungus. United States Patent No. 6,137,032, issued Oct. 24, 2000

Cheng K-J, Selinger LB, Liu J-H, Hu Y, Forsberg CW and **Moloney MM** (1999) Xylanase obtained from an anaerobic fungus. United States Patent No. 5,948,667, issued Sept. 7, 1999.

Moloney MM, van Rooijen GJH, and Boothe J (1999) Oil bodies and associated proteins as affinity matrices. United States Patent No. 5,856,452, issued Jan. 5, 1999

Moloney MM (1999) Preparation of heterologous proteins on oil bodies. United States Patent No. 5,948,682, issued Sept. 7, 1999.

Deckers HM, van Rooijen GJH, Boothe J, Goll J, **Moloney MM** and Mahmoud S (1998) Uses of Oil bodies. PCT Patent Application WO 98/53698. Published 12/98

Moloney MM, Alcantara J and van Rooijen GJH (1998) Method for cleavage of fusion proteins. PCT Application WO 98/49326. Published 11/98

Moloney, MM (1998) Transformation and foreign gene expression in *Brassica spp.* United States Patent. No. 5,750,871, issued May 12, 1998.

Moloney MM, van Rooijen GJH, and Boothe J (1998) Oil bodies and associated proteins as affinity matrices. PCT Application WO 98/27115. Published 6/98

Moloney MM, Boothe J, and van Rooijen GJH (1996) Oil bodies and associated proteins as affinity matrices. United States Patent No. 5,856,452, issued Jan. 5/99

Moloney, MM (1995) Transformation and foreign gene expression in *Brassica spp.* United States Patent No. 5,463,174, issued Oct. 31, 1995.

Moloney, MM (1993) Transformation and foreign gene expression in *Brassica spp.* United States Patent No. 5,188,958, issued Feb. 23, 1993.

Moloney MM (1992) Oil body protein cis elements as regulatory signals for seed specific expression. This patent describes the cis-acting sequences of oleosin genes which may be used to drive seed specific expression in transgenic plants. United States Patent No. 5,792,922, issued November 8, 1998

Moloney MM (1991) Oil body proteins as carriers of high value proteins. This patent describes a novel procedure for the production of high value peptides such as pharmaceuticals, enzymes, peptide hormones and adhesives in oilseeds. The method provides for a unique purification process which renders economical the production of such proteins in plant seeds. United States Patent No. 5,650,554, issued Jul 22/97

Hilliard J, **Moloney MM** (1989) Probe for electrofusion, electroporation or like procedure II. United States Patent No. 4,882,281, issued Nov. 21, 1989. This patent describes a novel probe with multiple electrodes for the performance of electroporation / electrofusion experiments. Permits the efficient transformation of bacterial or plant cells using a voltage generator of less than 600V.

Moloney, MM and Hilliard, J (1986) A probe for electrofusion, electroporation or like procedure. US Patent No. 4,695,547. This patent describes a novel probe and set-up for the performance of electroporation/electrofusion experiments. The patent has been licensed by Hoefer Scientific of San Fransisco CA and is the basis of their 'Progenitor' series of electroporation devices.

Moloney, MM (1986) Transformation of *Brassica spp* using Agrobacterium vectors. Applied U.S. Patent Office, May 1986. European claims accepted May 1987 #868,640. This patent describes a route to the production of transgenic *Brassica* plants using Agrobacterium vectors. Claims involve several target tissues. Issued March/92.

Impact and Contributions:

Evidence for the impact of my work in plant molecular biology and biotechnology can be found in several examples. I served for 3 years (1995-1999) as Editor of The Plant Journal, the second most cited journal in plant biology. I am a member of the advisory board at the Plant Biotechnology Institute in Saskatoon. I have several consulting relationships with both government departments and private industry. I act as a consultant for Dow AgroSciences Canada and frequently for the Provincial and Federal Government. I have served on NSERC (Strategic 1992-1995) and Alberta Government grant selection committees. I am an invited speaker to many international conferences including the TIGR International Genomics Conference, 2000, those of the International Society for Plant Molecular Biology (Amsterdam, 1994; Singapore, 1997), International Society for Fats and Oils Research (The Hague, 1995) and The Biochemical Society (Bristol, U.K., 1996). I give 7-10 invited seminars per year at Canadian, European or U.S. Universities. I was also the Chair and Co-organizer of the International Society for Plant Molecular Biology Congress in Quebec City, June, 2000.

The award of an NSERC Industrial Research Chair to me in 1995, underlines the interest and commitment of industry to our work and attests to our desire to convert, wherever possible, basic discovery research into useful technology. I was the winner of the Alberta Science and Technology foundation (ASTech) award for "outstanding leadership in Alberta technology" in October of 1997. In 2002, I was appointed to the NSERC Council the governing body of the Natural Sciences and Engineering Research Council of Canada. I am the chair of the NSERC Committee on Research Partnerships which accounts for approximately 35% of NSERC's annual research budget and I am a member of the NSERC Executive Committee.

I am the founder and Chief Scientific Officer of SemBioSys Genetics Inc. a biotechnology company based in Calgary, Alberta. SemBioSys, founded in 1994, employs approximately 45 people and is one of the largest Canadian plant biotechnology companies. SemBioSys Genetics Inc. is a world leader in the expression and

manufacture of recombinant proteins for medicinal and industrial applications, using oilseed plants as the host for their production.

Significant Research Contributions:

Subcellular targeting and topology of lipophilic proteins

Abell, B.M., High, S., Moloney, M.M. (2002) Membrane Protein Topology of Oleosin is Constrained by its Long Hydrophobic Domain. Journal of Biological Chemistry: in press. This paper provides a fundamental analysis to targeting and topology of oleosins as examples of the most lipophilic proteins in nature. From this work it is clear that the length of a hydrophobic stretch is more important than its actual sequence in determining topology on a membrane. Furthermore, we demonstrated that novel signal-anchoring sequences can be derived from oleosins to display proteins on the cytoplasmic side of the ER.

van Rooijen GJH, Moloney MM (1995) Structural requirements of oleosin domains for subcellular targeting to the oil body. *Plant Physiol.*, **109**:13553-1361.

Abell BM, Holbrook LA, Abenes M., Murphy DJ, Hills MJ, Moloney MM (1997) Role of the Proline Knot Motif in Oleosin Endoplasmic Reticulum Topology and Oil Body Targeting. *Plant Cell*, **9**:1481-1493.

Oleosins represent a unique class of proteins throughout nature. They have the longest hydrophobic stretches of any protein studied to date. They undergo targeting to oil-bodies through a co-translational process and yet they do not enter the secretion pathway. They associate with a translocation pore on the ER, but do not undergo cleavage of a signal sequence. Our work has provided the basis for understanding which parts of the oleosin are essential to correct subcellular trafficking and how the topology of the protein is maintained on oil-bodies. This work had broader implications for all lipophilic proteins including those of non-plant origin.

Use of oleosins as carriers for recombinant proteins

van Rooijen GJH, Moloney MM (1995) Plant seed oil-bodies as carriers for foreign proteins. *Biotechnology*, **13**:72-77.

Parmenter DL, Boothe JG, van Rooijen GJH, Yeung EC, Moloney MM (1995) Production of biologically active hirudin in plant seeds using oleosin partitioning. *Plant Mol. Biol.* **29**:1167-1180.

Liu J-H, Selinger LB, Cheng K-J, Beauchemin, KA, Moloney, MM (1997) Plant seed oil-bodies as an immobilization matrix for a recombinant xylanase from the rumen fungus *Neocallimastix patriciarum. Molecular Breeding* **3**:463-470

Moloney, Maurice M. **US Patent 5,650,554** "Oil Body Proteins as Carriers of High-Value Peptides in Plants" Issued Jul. 22, 1997

As a corollary to our work on oleosin targeting, we discovered that oleosin fusion proteins were also capable of efficient targeting to oil-bodies. This led to the hypothesis that recombinant proteins could be attached to oil-bodies in transgenic plants and then separated from other cellular contents based on floatation centrifugation. This liquid-liquid separation could be performed inexpensively and result in a versatile platform for producing recombinant proteins in seeds. In the above-cited papers this hypothesis was validated and several examples of recombinant proteins have been produced this way.

The technology has been patented and now comprises several patent families worldwide. These patents were used to form SemBioSys Genetics Inc., a Calgary-based biotechnology company currently employing 35 staff.

Regulation of seed-specific gene expression

Plant AL, van Rooijen GJH, Anderson CP, Moloney, MM (1994) Regulation of an *Arabidopsis* oleosin gene promoter in transgenic *Brassica napus*. *Plant Mol. Biol.* **25**:193-205.

Crowe AJ, Abenes M, Plant A, Moloney MM (2000) The seed-specific transactivator, AB13, induces oleosin gene expression. *Plant Science* **151**:171-181.

Abenes M, Holbrook L, Moloney, MM (1997) Transient expression and oil body targeting of an Arabidopsis oleosin-GUS reporter fusion protein in a range of oilseed embryos. *Plant Cell Rep.* **17**:1-7.

Moloney, Maurice M. **US Patent 5,792,922** "Oil-Body Proteins Cis-Elements as Regulatory Signals" Issued Aug. 11, 1998.

My laboratory has performed a substantial amount of work on seed-specific gene regulation. We have been particularly interested in the modulation of transcriptional activity by plant hormones especially abscisic acid. My lab has performed the characterization of oleosin gene promoters and their cis-elements and their interaction with key

transcription factors such as ABI3. This work has resulted in greater insights into the factors needed for high level seed-specific expression, and in the isolation, use and patenting of a broad family of seed-specific promoters with uses in modification of lipid, starch or protein deposition in developing seeds.

Exhibit B

8 Webster, M.E. (1968) Fed. Proc. 27, 84—89
9 Haberland, G.L. (1973) New Aspects of Trasylol Therapy (Haberland, G.L. and Lewis, D.H. eds), pp. 1—7, F.J. Schattauer Verlag, Stuttgart, New York
10 Vogelaar, E.F., van der Meer, J., Brummelhuis, H.G.J. and Krijnen, H.W. (1972) Protides Biol. Fluids 20, 447—450
11 Cohn, E.J., Strong, L.E., Hughes, W.L., Mulford, D.J., Ashworth, J.N., Melin, M. and Taylor, H.J. (1946) J. Am. Chem. Soc. 68, 459—475
12 Harowitz, H.J., Wilcox, W.P. and Fujimoto, M.M. (1963) Blood 22, 35—43
13 Geratz, J.D. (1970) Thromb. Diath. Haemorrh. 23, 486—499
14 Press, E.M., Porter, R.R., Cebra, J. (1960) Biochem. J. 74, 501—514
15 Webster, M.E. and Prado, E.S. (1970) Methods in Enzymology (Perlmann, G.E. and Lorand, L., eds.), Vol. 19, pp. 681—699, Academic Press, New York
16 Ouchterlony, O. (1953) Acta Pathol. Microbiol. Scand. 32, 231—240
17 Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265—275
18 Schultze, H.E. and Heremans, J.F. (1966) Molecular Biology of human proteins, Vol. I, pp. 250—252, Elsevier, New York
19 Colman, R.W., Mettler, L. and Sherry, S. (1969) J. Clin. Invest. 48, 11—12
20 Colman, R.W., Mettler, L. and Sherry, S. (1969) J. Clin. Invest. 48, 23—32
21 Webster, M.E. and Gilmore, J.P. (1965) Biochem. Pharmacol. 14, 1161—1163
22 Sherry, S., Alkjaersig, N.K., Fletcher, A.P. (1966) Throm. Diath. Haemorrh. 16 (Suppl. 20), 243—251
23 Avrameas, T. and Ternynck, T. (1969) Immunochemistry 6, 53—66
24 Hial, V., Diniz, C.R. and Mares-Guia, M. (1974) Biochemistry 13, 4311—4318
25 Dixon, M. (1953) Biochem. J. 55, 170—171
26 Sampaio, C., Wong, S. and Shaw, E. (1974) Arch. Biochem. Biophys. 165, 133—139
27 Davies, G.E. and Lowe, J.S. (1970) Adv. Exp. Med. Biol. 8, 453—460
28 Geratz, J.D. and Webster, W.P. (1971) Arch. Int. Pharmacodyn. 194, 359—370
29 Harpel, P.C. (1971) J. Clin. Invest. 50, 2084—2090
30 Schrieber, A.D., Kaplan, A.P. and Austen, K.F. (1973) J. Clin. Invest. 52, 1394—1401
31 Bagdasarian, A., Lahiri, B., Talmo, R.C., Wong, P. and Colman, R.W. (1974) J. Clin. Invest. 54, 1444—1454
32 Dyce, P., Wong, T., Adham, N., Mehl, J. and Haverback (1967) Clin. Res. 15, 101
33 Webster, M.E. and Innerfield, I. (1965) Enzymol. Biol. Clin. 5, 129—148
34 Harpel, P.C. (1973) J. Exp. Med. 138, 508—521
35 Back, N., Wilkens, H. and Streger, R. (1968) Ann. N.Y. Acad. Sci. 146, 491—509
36 Webster, M.E. and Maling, H.M. (1970) Adv. Exp. Biol. Med. 8, 493—499
37 Webster, M.E. (1968) Fed. Proc. 27, 84—89
38 Kellermeyer, R.W. and Graham, R.C. (1968) New Eng. J. Med. 279, 859—866
39 Asghar, S.S., Pondman, K.W. and Cormane, R.H. (1973) Biochim. Biophys. Acta 317, 539—548
40 Schoenback, E.B. and Greenspan, E.M. (1948) Medicine 27, 327—377
41 Ivady, G., Paldy, L., Koltay, M., Toth, G. and Covacs, Z. (1967) Lancet I, 616—617
42 Ashley, J.N., Berg, S.S. and Lucas, J.M.S. (1960) Nature 185, 461
43 Ellis, H.V., Johnson, A.R. and Moran, N.C. (1970) J. Pharmacol. Exp. Ther. 175, 627—631
44 Goodman, L.S. and Gilman, A. (1970) The Pharmacological Basis of Therapeutics, 4th edn. pp. 1144—1153, The Mac Millan Co., New York
45 Wien, R. (1943) Ann. Trop. Med. Parasit. 37, 1—18

Biochimica et Biophysica Acta, 438 (1976) 265—272

BBA 67831

PEPTIDE SUBSTRATES FOR CHYMOSIN (RENNIN)

KINETIC STUDIES WITH PEPTIDES OF DIFFERENT CHAIN LENGTH INCLUDING PARTS OF THE SEQUENCE 101-112 OF BOVINE κ-CASEIN

S. VISSER [a], P.J. VAN ROOIJEN [a], CECILE SCHATTENKERK [b] and K.E.T. KERLING [b]

[a] *Netherlands Institute for Dairy Research (NIZO), Ede and* [b] *Gorlaeus Laboratories Department of Organic Chemistry, University of Leiden, Leiden (The Netherlands)*

(Received October 27th, 1975)

Summary

Kinetic parameters have been determined for the reaction between milk-clotting chymosin (EC 3.4.23.4) and a series of peptides (or their methyl esters) including the amino acid sequence around the enzyme-sensitive Phe(105)-Met (106) bond of bovine κ-casein. In particular, the influence of the substrate's chain length on the kinetic parameters has been studied. Evidence is presented that in the model peptides studied the sequence -Ser-Phe-Met-Ala- with a further residue added to either end (in casu Leu(103) or Ile(108)) is necessary to induce any cleavage by the enzyme. When both the Leu(103) and Ile(108) residues form part of the peptide chain, a marked improvement of the substrate properties is observed. It is suggested that prolyl residues on either side of the sensitive peptide bond form additional sites for secondary enzyme-substrate interactions.

Introduction

Chymosin (rennin, EC 3.4.23.4) is a milk-clotting enzyme, the precursor of which is secreted by the fourth stomach of the calf [1]. During the last few decades it has been established that the process of milk clotting is initiated by the specific proteolysis of so-called κ-casein by the enzyme [2]. This initial action on κ-casein is restricted to the cleavage of the peptide bond 105-106 between phenylalanine and methionine [3—5]. To gain information how this specific cleavage is connected with the structure of the substrate, we have undertaken a systematic study of the substrate specificity of chymosin. For this purpose we have utilized a series of synthetic peptides of different chain length includ-

ing parts of the sequence

-Pro-His-Leu-Ser-Phe-Met-Ala-Ile-Pro-Pro-Lys-Lys-
101 103 105 106 108 110 112

around the chymosin-labile Phe-Met bond of bovine κ-casein.

In previous communications [6,7] some of us have reported on the substrate properties of a number of such peptides. Conclusions were based on high-voltage paper electrophoretic patterns after prolonged incubation of the peptides with the enzyme. The main features arising from this work were: First, a minimum chain length of five amino acid residues seems to be a prerequisite for cleavage by the enzyme. Second, the sequence Leu(103)-Ser(104) is of importance. When this sequence is reversed, a case studied by Hill [8,9], a sharp decrease in the rate of cleavage is observed. Third, further chain extension at either side of the sensitive bond enhances the capacity of the peptide to function as a substrate for chymosin.

Independently of this work, other groups have also studied the action of chymosin on parts of κ-casein containing the Phe-Met bond in question. Hill [8, 9] concluded from his experiments with short synthetic peptides as well as with photo-oxidized κ-casein [10] that the enzymic action is accelerated by serine and histidine side chains located close to the Phe-Met bond. Polzhofer [11] found a synthetic pentadecapeptide to be split rapidly and calculated a Michaelis constant for this reaction. Furthermore, he concluded that the His (102) residue has an important role, since the hexapeptide His-Leu-Ser-Phe-Met-Ala appeared to be split by chymosin whereas the pentapeptide Leu-Ser-Phe-Met-Ala was found to be resistant to enzymatic cleavage. Evidence for the importance of the hydroxyl group of Ser(104) was presented by Raymond et al. [12,13] working with peptide substrates slightly different from the parent primary structure.

In the present paper kinetic studies are reported in which special attention has been given to the influence of the chain length of the substrate on the kinetics of its reaction with chymosin. This approach permits an investigation of the effect of secondary enzyme-substrate interactions and may lead to some conclusions about the size of the enzyme's active centre [14].

Materials and Methods

Peptides and their derivatives were synthesized and characterized as described by Schattenkerk et al. [15]. The specificity of their cleavage by chymosin was checked by thin-layer chromatography, paper electrophoresis and by N-terminal group analysis using dansyl chloride [16]. Crude chymosin, isolated as an extract ("rennet") from stomachs of newborn calves, was supplied by the "Coöperatieve Stremsel- en Kleurselfabriek" (Leeuwarden, The Netherlands). From this extract chymosin was purified by DEAE-Sephadex chromatography as described by De Koning [3,4]. The milk-clotting activity of the purified enzyme amounted to $6.5 \cdot 10^6$ Soxhlet units [17,18]. As a reference the "Netherland Standard for the determination of (calf-)rennet strength" furnished by the Government Dairy Station (Leiden, The Netherlands), was used.

Enzyme solution. The enzyme was dissolved to a final concentration of 0.418 M in 0.05 M sodium acetate buffer containing 1 M NaCl (pH 5.25) assuming a molecular weight of 30 000 [1,3,4]. Small portions of the enzyme solution, sufficient for a 1-day experiment, were kept frozen until needed in tightly closed 1 ml vials. With this procedure the proteolytic activity remained unaffected for many months.

Substrate solutions. Freeze-dried peptide (5—10 mg for one duplicate experiment) was brought into contact with a suitable volume of 0.05 M sodium acetate buffer (pH 4.7). After centrifugation of non-dissolved material *, the supernatant was utilized as a stock solution, the concentration of which was established as follows. Duplicate samples, to which known amounts of L-norleucine had been added as an internal standard, were made 6 M in HCl and heated at 110°C in small evacuated tubes for 22 h. The contents were then evaporated to dryness in vacuo. The residues were taken up with sodium citrate buffer (pH 2.2) and analyzed with the aid of an automatic amino acid analyzer (JEOL JLC-5AH). The concentration of peptide initially present was calculated by comparison with a standard mixture of amino acids analyzed in the same way.

Duplicate series of 5—8 different concentrations were prepared by diluting samples of the stock solution with 0.05 M sodium acetate (pH 4.7).

Kinetic measurements. Reactions were carried out at 30°C in a 0.05 M sodium acetate buffer, pH 4.7. The enzymatic cleavage was monitored by the automated ninhydrin assay described elsewhere [20]. Each experiment was started with equal volumes of substrate solution (800 μl) in standard size reaction tubes (5.5 × 0.6 cm). The time of mixing the enzyme with the substrate was kept as short as was necessary to ensure complete mixing (5 s). By doing so, we could minimize the effect of a gradually decreasing initial reaction rate which in preliminary experiments had been observed when the reaction mixture was continuously stirred. This effect, the magnitude of which also depended on the total volume stirred, is probably to be attributed to an inactivation of the enzyme at the air-water interface. The same phenomenon was observed when a substrate in which methionine-106 had been replaced by a different residue (e.g. leucine or norleucine), was used. This rules out air oxidation of the substrate's methionine side chain, resulting in decreasing substrate properties (cf. Results), as the main source of this effect.

Evaluation of kinetic parameters. The kinetic parameters, V and K_m, were calculated from the collected data of two independent experiments each carried out with 5—8 substrate concentrations. A BASIC-programmed Hewlett-Packard calculator, model 9830 A, was used routinely to determine slope and intercept from plots of c/v vs c [21] and $1/v$ vs $1/c$ [22]. The programme provides all necessary calculations including a weighted (v^{-4}) least-squares fit to a straight line with standard errors for the computed parameters [23] and the correction for hydrolytic cleavage discussed below.

Correction of parameters for the extent of initial hydrolysis. For the process:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P_1 + P_2,$$

* Addition of an organic solvent, which is sometimes used to enhance the solubility of substrates, has been avoided in this study as this may affect the kinetics of enzymic action [19].

(E = enzyme, S = substrate, and P = product) the rate equation can be written as:

$$v_0 = \frac{Vc_s^0}{K_m + c_s^0}, \quad (1)$$

where v_0 is the initial velocity; c_s^0 the initial substrate concentration; V the maximal velocity (= k_{cat} × total enzyme concentration) and K_m the Michaelis constant (= $(k_{cat} + k_{-1})/k_1$). When the substrate concentration changes to a considerable extent during the time course of measurement (t) the integrated rate equation should be used:

$$Vt = c_p^t + K_m \ln \frac{c_s^0}{c_s^0 - c_p^t} \quad (2)$$

where c_p^t stands for the concentration of product formed. After expanding the logarithm and rearrangement, one obtains

$$\frac{c_p^t}{t} + \frac{K_m\left\{\frac{1}{2}\left(\frac{c_p^t}{c_s^0}\right)^2 + \frac{1}{3}\left(\frac{c_p^t}{c_s^0}\right)^3 + \frac{1}{4}\left(\frac{c_p^t}{c_s^0}\right)^4 + \ldots\right\}}{t\left(\frac{c_s^0 + K_m}{c_s^0}\right)} = \frac{Vc_s^0}{K_m + c_s^0} \quad (3)$$

The ratio c_p^t/c_s^0 is a measure of the extent of hydrolysis. Its value can be established for each substrate concentration [20]. In Eq. 3 the second term on the left-hand side represents the correction of the apparent initial velocity c_p^t/t for substrate depletion. Since this term contains the parameter K_m, we have used an iterative procedure for the evaluation of correct Michaelis parameters. As a starting value of K_m the result from a linear plot with the apparent initial velocities was taken.

Results

Identical kinetic parameters, V and K_m, were obtained by analysing the c_s^0/v vs c_s^0 plot and the $1/v$ vs $1/c_s^0$ plot only after proper weighting for inversion [24].

In Table I the kinetic parameters k_{cat}, K_m and k_{cat}/K_m for a number of peptides are presented together with their standard errors. The experimental conditions are specified in columns 4 and 5. The maximal depletion of substrate during the time used for the determination of initial velocities, is given in column 6. In the last column the number of experimentally determined initial velocities over the concentration range given, is listed. As may be expected, the values of the separate parameters, k_{cat} and K_m, are more reliable when K_m is well within the range of substrate concentrations.

The pentapeptide ester Ser-Phe-Met-Ala-Ile-OMe (I), which previously [6] was found to be split to a reasonable extent during a 24-h incubation period with the enzyme, appeared to be a poor substrate in terms of k_{cat}/K_m as compared with the substrates V—XIII. The same can be said of other peptides with N-terminal serine and chains extended in the C-terminal direction (substrates

TABLE I

KINETIC PARAMETERS OF THE REACTION BETWEEN CHYMOSIN AND SYNTHETIC SUBSTRATES CONTAINING PARTS OF THE κ-CASEIN SEQUENCE

Pro - His - Leu - Ser - Phe - Met - Ala - Ile - Pro - Pro - Lys - Lys -
101 103 105 106 108 110 112

All experiments were carried out at 30°C in 0.05 M sodium acetate buffer (pH 4.7).

Substrate	Number of residues	Substrate concentration (mM)	Enzyme[a] concentration (nM)	Percentage of hydrolysis	k_{cat}[a] (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})	n
Ser(104)—Ile(108)OMe	5	0.43—1.72	2960	0.8— 1.2	0.33 ± 0.10[b]	8.5 ± 2.7[b]	0.038 ± 0.002	10
Ser(104)—Pro(109)OMe	6	0.20—1.60	1490	0.9— 3.2	1.05 ± 0.45[b]	9.2 ± 4.0[b]	0.114 ± 0.007	10
Ser(104)—Pro(110)OMe	7	0.66—1.76	775	2.1— 3.3	1.57 ± 0.82[b]	6.8 ± 3.6[b]	0.231 ± 0.022	12
Ser(104)—Lys(111)OH	8	0.30—1.20	1540	4.4— 5.4	0.75 ± 0.15[b]	3.2 ± 0.6[b]	0.239 ± 0.013	10
Leu(103)—Ile(108)OMe	6	0.10—0.80	66	8.6—14.6	18.3 ± 0.9	0.85 ± 0.05	21.6 ± 0.7	15
Leu(103)—Pro(109)OMe	7	0.08—0.64 0.10—0.83	38 30	12.2—21.3 9.1—17.3	37.5 ± 1.7 38.7 ± 1.4	0.71 ± 0.04 0.67 ± 0.03	52.8 ± 1.6 57.5 ± 1.7	16 13
Leu(103)—Pro(110)OMe	8	0.09—0.76	19	7.9—18.3	43.3 ± 2.3	0.41 ± 0.03	105.1 ± 6.9	15
Leu(103)—Lys(111)OH	9	0.06—0.50	19	7.4—14.2	33.6 ± 1.2	0.43 ± 0.02	78.3 ± 2.3	13
Leu(103)—Lys(112)OH[c]	10	0.11—0.45 0.06—0.45	43 29	15.9—24.9 10.6—18.0	31.4 ± 1.6 29.0 ± 1.2	0.49 ± 0.03 0.43 ± 0.02	63.7 ± 2.2 66.9 ± 2.1	13 16
Leu(103)—Lys(112)OH[d]	10	0.05—0.37	29	10.9—19.2	25.3 ± 1.5	0.40 ± 0.03	63.0 ± 2.6	7[f]
Leu(103)—Lys(112)OH[e] [(106)Met(O)]	10	0.09—0.36	143	9.7—12.3	7.1 ± 0.4	0.84 ± 0.05	8.41 ± 0.16	6[f]
His(102)—Ile(108)OMe	7	0.09—0.65	52	8.8—17.5	16.0 ± 0.8	0.62 ± 0.03	30.8 ± 1.4	12
Pro(101)—Ile(108)OMe[c]	8	0.10—0.83 0.09—0.71	15 14	5.4—13.0 6.2—16.9	34.8 ± 0.5 32.3 ± 0.5	0.37 ± 0.01 0.31 ± 0.01	94.6 ± 2.1 105.9 ± 2.6	15 16

[a] Calculated assuming a molecular weight of 30 000 for the enzyme.
[b] Values of the separate parameters, k_{cat} and K_m, are rather uncertain, since the v vs c_s^0 plot largely showed first-order kinetics.
[c] Results are given of two duplicate experiments done at a time interval of at least one month.
[d] After preliminary treatment with 2-mercaptoethanol.
[e] After preliminary oxidation of the methionyl residue to its sulphoxide according to the method described by Iselin [25].
[f] Data obtained from a single experiment.

II—IV). Addition of a leucyl residue to the N-terminal part of the peptides I–IV caused an increase of the k_{cat}/K_m parameter by more than two orders of magnitude (substrates V—VIII). Extension with a histidyl residue at the N-terminal side of the hexapeptide ester V hardly influenced the k_{cat}; the increase of the k_{cat}/K_m ratio was predominantly brought about by a change of K_m (XII). Further extension in N-terminal direction with a prolyl residue brought about an increase in k_{cat} and a decrease in K_m. The k_{cat} value (32—35 s^{-1}) found for this octapeptide ester XIII is about the same as that of the heptapeptide ester VI obtained by extension with only one (prolyl) residue at the C-terminal side of peptide V. Comparing the kinetic parameters of the substrates V and VI, one sees that for the longer peptide the K_m is somewhat lower while the k_{cat} is doubled. Addition of a second proline to this sequence, leading to the octapeptide ester Leu-Ser-Phe-Met-Ala-Ile-Pro-Pro-OMe (VII), resulted in a further decrease of K_m leaving k_{cat} almost unchanged. As soon as the lysine-111 had been joined to the sequence, the k_{cat} value decreased slightly while the K_m remained unaltered (VIII). However, no definite conclusions in terms of electrostatical effects can be drawn from this, since a free carboxylic end group was introduced simultaneously in the peptide chain in place of the methyl ester group, leaving the net peptide charge unaffected. In fact, an extra positive charge was added only by introduction of the second lysyl residue at position 112 (IX). This hardly altered the kinetic parameters as compared with those of the substrate VIII.

Treatment of the peptide IX with 2-mercaptoethanol to eliminate the possible effect of methionine oxidation did not influence the enzyme kinetics (X). Apparently methionine sulphoxide was not present in the peptide preparation in significant amount. This argument is corroborated by the fact that careful oxidation of the methionyl residue to the sulphoxide resulted in a striking fall of the substrate properties as expressed by a decreasing k_{cat} and an increasing K_m (XI).

Discussion

Considering the results listed in Table I, it appears that the leucyl residue at position 103 is of great importance for the rate of hydrolysis of the substrate, as revealed by a sharp increase in k_{cat} after the introduction of this residue. Two factors might be responsible for this effect. First, the extension of the peptide backbone at the "left-hand" side, which, in addition, may lead to a suitable location of the N-terminal charge with respect to a counterion in the enzyme. Second, the increase in hydrophobicity of the peptide by the addition of the leucyl side chain.

From earlier work [7] it follows that the isoleucine-108 also strongly influences the rate of hydrolysis, since the peptide ester Leu-Ser-Phe-Met-Ala-OMe (thus containing Leu(103) but missing Ile(108)) showed equally poor substrate properties as the peptide ester Ser-Phe-Met-Ala-Ile-OMe (I) listed in Table I. Further evidence for an important role of the residue in position 108 was presented by Raymond et al. [12,13] in comparing the peptides Leu-Ser-Phe(NO_2)-Nle-Ala-OMe (k_{cat}/K_m = 0.1 mM^{-1} s^{-1}) and Leu-Ser-Phe(NO_2)-Nle-Ala-Leu-OMe (k_{cat}/K_m = 12.7 mM^{-1} s^{-1}) as substrates for chymosin.

The tetrapeptide esters Leu-Ser-Phe-Met-OMe, Ser-Phe-Met-Ala-OMe and Phe-Met-Ala-Ile-OMe were all found to be completely resistant to cleavage by chymosin [6].

From the above it is apparent that a minimum chain length of five amino acid residues including the sequence -Ser-Phe-Met-Ala- is essential to bring about any cleavage of the Phe-Met bond. A large jump in the substrate quality is effected when both the Leu(103) and the Ile(108) form part of the peptide chain, as is reflected by the marked differences in k_{cat}/K_m between the substrate series I—IV and series V—VIII. The contribution to the rate of proteolysis of substrate groups at some distance from the bond to be split by the enzyme has also been reported for the hydrolysis of peptides by other proteases such as pepsin [26,27], and chymosin-like enzymes from *Mucor miehei* [28] or from *Mucor pusillus* [29].

As is evident from Table I, additional sites of secondary interaction can be located on amino acid side chains more distant from the labile Phe-Met bond. These substrate groups, however, contribute to a much lesser extent to the substrate quality of the peptide than do the leucine-103 and isoleucine-108 discussed above. The kinetic parameters of the substrates V—XIII (except the oxidized peptide XI) all fall within the same range: k_{cat} 16—43 s^{-1} and K_m 0.3—0.9 mM. The most suitable substrates found in the present study were the octapeptide esters VII and XIII ($k_{cat}/K_m \approx 100$ s^{-1} mM^{-1}). These substrates contain, in addition to the important Leu(103) and Ile(108), one or two prolyl residues which evidently further add to the secondary enzyme-substrate interaction. The prolyl residues might also impart some stabilization to a preferential conformation of the substrate molecule.

A discussion about the role of the histidyl and lysyl side chains is hampered by the fact that all our experiments were carried out at the same pH and ionic strength. The pH of 4.7 has been chosen as being the optimum pH for the action of chymosin on small peptide substrates [12]. It cannot be ruled out that the substitution of an additional charged group in a peptide substrate will have a considerable effect on the pH optimum for enzymic cleavage. The conclusion that the introduction of the lysyl residues at the positions 111 and 112 would not greatly affect the substrate quality (cf. Table I) must therefore be considered with care. In addition to this, a second aspect has to be taken into account when considering the function of the side chain of histidine-102. One can conclude from Table I that under the conditions of our studies, coupling of a histidyl residue to the peptide ester V, leading to substrate XII, only influences the K_m of the reaction to a reasonable extent leaving the k_{cat} almost unchanged. This does not imply, however, that in a more extended molecule (e.g. in intact κ-casein) the positive charge of the histidine side chain will not have any effect on the k_{cat} of the enzymic cleavage. In the present case of small peptide esters the function of the protonated imidazole group may have been taken over completely or partly in the hexapeptide ester V by the positively charged N-terminal residue. Definite conclusions as to the function of the histidyl and lysyl side chains in the enzyme-substrate interactions have to await the results of further experimental work.

A marked influence on the kinetic parameters was observed when the methionyl residue of peptide IX had been oxidized to its sulphoxide (XI). This treat-

ment apparently affected the primary interactions between the enzyme and its substrate.

Acknowledgements

We are indebted to Professor Dr. E. Havinga (University of Leiden) and to Dr. P.J. de Koning (NIZO) for continuous interest and valuable discussions. We wish to thank Dr. H.J. Vreeman (NIZO) for critical remarks and mathematical guidance.

References

1 Foltmann, B. (1966) C. R. Trav. Lab. Carlsberg, 35, 143—231
2 MacKinlay, A.G. and Wake, R.G. (1971) Milk Proteins (McKenzie, H.A., ed.), Vol. II, pp. 175—215, Academic Press, New York
3 De Koning, P.J. (1967) Thesis, University of Amsterdam
4 De Koning, P.J. (1968) Neth. Milk Dairy J. 22, 121—124
5 Jollès, J., Alais, C. and Jollès, P. (1968) Biochim. Biophys. Acta 168, 591—593
6 Schattenkerk, C., Holtkamp, I., Hessing, J.G.M., Kerling, K.E.T. and Havinga, E. (1971) Recl. Trav. Chim. Pays-Bas 90, 1320—1322
7 Schattenkerk, C. and Kerling, K.E.T. (1973) Neth. Milk Dairy J. 27, 286—287
8 Hill, R.D. (1968) Biochem. Biophys. Res. Commun. 33, 659—663
9 Hill, R.D. (1969) J. Dairy Res. 36, 409—415
10 Hill, R.D. and Laing, R.R. (1965) J. Dairy Res. 32, 193—201
11 Polzhofer, K.P. (1972) Tetrahedron 28, 855—865
12 Raymond, M.N., Garnier, J., Bricas, E., Cilianu, S., Blasnic, M., Chaix, A. and Lefrancier, P. (1972) Biochimie 54, 145—154
13 Raymond, M.N., Bricas, E. and Mercier, J.C. (1973) Neth. Milk Dairy J. 27, 298—303
14 Schechter, I. and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157—162
15 Schattenkerk, C., Voskuyl-Holtkamp, I. and Bokhorst, R. (1973) Recl. Trav. Chim. Pays-Bas 92, 92—116
16 Gray, W.R. (1972) Methods Enzymol. 25b, 121—138
17 Soxhlet, F. (1877) Milchzeitung 6, 497—501
18 Soxhlet, F. (1877) Milchzeitung 6, 513—514
19 Hollands, T.R. and Fruton, J.S. (1968) Biochemistry 7, 2045—2053
20 Vreeman, H.J., van Rooijen, P.J. and Visser, S. submitted for publication
21 Hanes, C.S. (1932) Biochem. J. 26, 1406—1421
22 Lineweaver, H. and Burk, D. (1934) J. Am. Chem. Soc. 56, 658—666
23 Cleland, W.W. (1967) Adv. Enzymol. 29, 1—32
24 Dowd, J.E. and Riggs, D.S. (1965) J. Biol. Chem. 240, 863—869
25 Iselin, B. (1961) Helv. Chim. Acta 44, 61—78
26 Fruton, J.S. (1970) Adv. Enzymol. 33, 401—443
27 Fruton, J.S. (1974) Acc. Chem. Res. 7, 241—246
28 Oka, T. and Morihara, K. (1973) Arch. Biochem. Biophys. 156, 552—559
29 Oka, T., Ishino, K., Tsuzuki, H., Morihara, K. and Arima, K. (1973) Agric. Biol. Chem. (Tokyo) 37, 1177—1184

Biochimica et Biophysica Acta, 438 (1976) 273—286

BBA 67816

ACTIVATION OF FIBROBLAST PROCOLLAGENASE BY MAST CELL PROTEASES

HENNING BIRKEDAL-HANSEN *, CHARLES M. COBB, ROBERT E. TAYLOR and HAROLD M. FULLMER **

Institute of Dental Research, School of Dentistry, University of Alabama in Birmingham, Birmingham, Ala. 35294 (U.S.A.)

(Received December 4th, 1975)

Summary

Proteases capable of activating procollagenase from gingiva and from fibroblast and macrophage monolayer cultures were harvested from homogenates of canine tumor mast cells. The mast cell proteases lysed casein and Azocoll but not native collagen. In low salt concentrations the enzymes existed as high molecular weight complexes, which were dissociated by increasing the salt concentration above 1.0 M (NaCl, KCl). Gel filtration in 1.4 M KCl separated the protease activity into three peaks, all of which activated procollagenase. Two of the enzymes showed substrate specificities (hydrolysis of *p*-tosyl-L-arginine methyl ester and benzoyl-tyrosine ethyl ester) and reactive center reactivities similar to pancreatic trypsin and chymotrypsin. Based on gel filtration, apparent molecular weights of 160 000 (*p*-tosyl-L-arginine methyl ester esterase), 90 000 (main procollagenase activator) and 36 000 benzoyl-tyrosine ethyl ester esterase) were determined. Activation of procollagenase resulted in a 18—20 000 decrease of the molecular weight. The activation was directly related to the amount of activator added within certain limits. Further addition of activator resulted in proteolytic inactivation of collagenase.

Introduction

Collagenases (EC 3.4.24.3) are enzymes capable of cleaving the helical portion of native collagen molecules in a characteristic manner. A series of studies have shown that animal collagenase may exist in a latent form [1—6]. More-

* Present address: Institute for Cariology and Endodontics, Royal Dental College, 160 Jagtvej, DK-2100 Copenhagen O, Denmark.
** Requests for reprints should be addressed to Dr. Harold M. Fullmer at the Institute of Dental Research, University of Alabama in Birmingham, Birmingham, Ala. 35294, U.S.A.

Exhibit C

STUDIES ON POLYPEPTIDES. XIII[1]

Relation between structure and capacity to function as rennin substrate.

(Short communication)

BY

C. SCHATTENKERK, I. HOLTKAMP, J. G. M. HESSING, K. E. T. KERLING and E. HAVINGA

(Laboratory of Organic Chemistry, The University, P.O. Box 75, Leiden)

With the aim of exploring the substrate-specificity of the enzyme rennin, which triggers the milk clotting process by specifically splitting the phenylalanine-methionine linkage in κ-casein[2], we synthesized a series of free peptides or their methyl esters containing the Phe-Met bond and increasing parts of the sequence adjacent to this dipeptide at both sides*. The susceptibility of the peptides towards hydrolysis by rennin was tested by incubation at 37° for 24 hours with a solution of rennin** in 0.05 *M*-sodium acetate buffer (pH 4.7).

The hydrolysis of the substrates was checked by paper electrophoresis at pH 1.8; the substrates and their possible splitting products were employed as references.

The enzyme-substrate ratio was 1/1500; the specific activity of the enzyme preparation amounted to 6.62×10^5 units/mg protein. The results represented in Fig. 1, give a clear picture of the capacity of the peptides to function as a substrate for rennin. This capacity was found to increase with increasing length of the peptide chain containing the Phe/Met bond. A high rate of reaction is observed with the hexapeptide Leu-Ser-Phe-Met-Ala-Ile; shortening the chain from either side diminishes the rate of reaction. We compared the peptides containing the sequence Ser-Leu-Phe-Met-Ala[3] with the sequence Leu-Ser-Phe-Met-Ala[2]. Under identical conditions the latter proved to be the better substrate.

It is intended to collect more detailed information about the features of the peptides essential to functioning as substrates for rennin, by enzyme-kinetic studies, varying a) the reaction time and b) enzyme-substrate ratio.

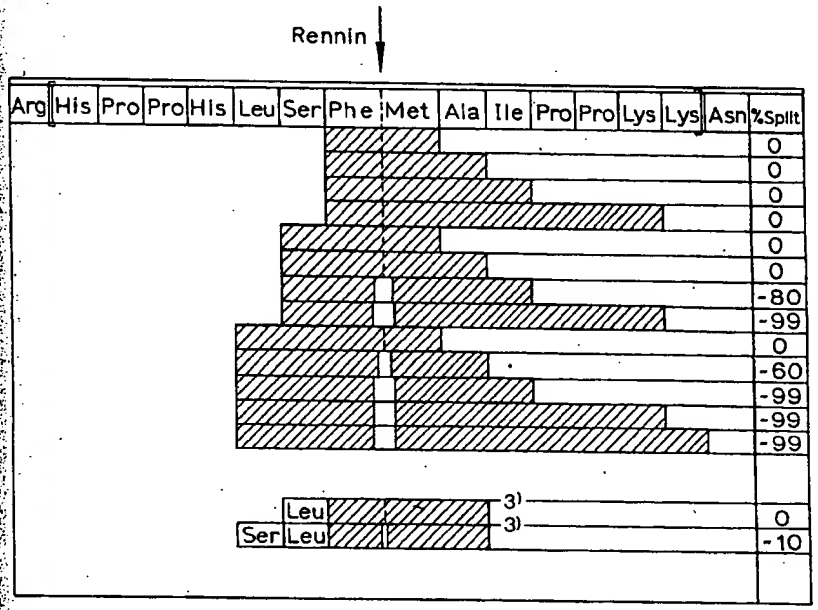


Fig. 1. Action of rennin on synthetic peptides containing the amino acid sequence around the Phe-Met bond of κ-casein.

* The syntheses and the physical properties of the peptides, represented in Fig. 1, will be dealt with in a forthcoming paper. Generally, for the enzymic experiments the methyl esters were used on account of their satisfactory solubilities. For the peptides containing lysine this proved not necessary, the solubilities of the free peptides being sufficiently high.

** Rennin preparations were kindly put at our disposal by Dr. *P. J. de Koning* of the N.I.Z.O. (Netherland Institute of Dairy Research), whom we wish to thank for his cooperation and interest.

[1] Part XII. *J. S. de Graaf, A. C. A. Jansen, C. Schattenkerk, K. E. T. Kerling* and *E. Havinga*, Rec. Trav. Chim. **90**, 301 (1971).

[2] See *e.g. J. Jollès, C. Alais, P. Jollès*, Bioch. Biophys. Acta. 168, 591 (1968); *P. J. de Koning*, Thesis University of Amsterdam 1967. Verslag no. 104 NIZO, Ede 1967.

[3] Cf. *R. D. Hill*, Bioch. and Biophys. Research Comm. 33, 659 (1968); J. Dairy Research 36, 409 (1969).

Acknowledgements

The authors wish to mention the valuable contributions that Miss *R. Bokhorst* and Messrs. *H. G. van Faassen* and *G. M. A. van Beynum* made to parts of the experimental work. The investigations were supported (in part) by the Netherlands Foundation of Chemical Research (S.O.N.) with financial aid from The Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

(*Received January 7th*, 1971)

A NOTE ON THE ROSENMUND REDUCTION OF ACID CHLORIDES

(Short Communication)

BY

J. A. PETERS and H. VAN BEKKUM

(Laboratory of Organic Chemistry, Delft University of Technology, Julianalaan 136, Delft, The Netherlands)

According to a recent announcement[1] acid chlorides can be hydrogenated to aldehydes over palladium at ambient temperature and at a pressure of about 3 atm, when fused sodium acetate is added to the solvent (xylene). This prompts us to report our results in the field of acid chloride reduction.

So far, the procedure commonly used in the catalytic reduction of acid chlorides was the original method of *Rosenmund*[2,3] which involves boiling toluene or xylene as the solvent and sulfur-regulated palladium on barium sulfate as the catalyst. The more convenient method of *Sakurai* and *Tanabe*[4] using *N,N*-dimethylaniline as hydrogen chloride acceptor and acetone as solvent has received almost no attention. With this modification acid chlorides can be reduced on palladium at room temperature and at atmospheric pressure. However, we observed that *N,N*-dimethylaniline itself is hydrogenated under these conditions which obscures the end-point of the reduction.

We have developed a modified *Sakurai-Tanabe* procedure for the synthesis of aliphatic and alicyclic aldehydes using palladium on carbon as the catalyst, ethyldiisopropylamine as the hydrogen chloride acceptor and acetone as the solvent. Ethyldiisopropylamine has, like *N,N*-dimethylaniline, the advantage of forming an acetone soluble hydrochloride, so the reaction mixture remains homogeneous. The present formulation enabled smooth reduction of aliphatic and alicyclic acid chlorides at room

[1] Chem. Eng. News [48], 41 (1970).
[2] *K. W. Rosenmund*, Chem. Ber. 51, 585 (1918).
[3] *E. Mosettig* and *R. Mozingo*, Org. Reactions 4, 362 (1948); *P. N. Rylander*, Catalytic Reduction over Platinum Metals, Academic Press, New York, 1967, p. 398.
[4] *Y. Sakurai* and *Y. Tanabe*, J. Pharm. Soc. Japan 64, 25 (1944) (Chem. Abstr. 45, 5613 (1951)).

Int. J. Peptide Protein Res. 47, 1996, 56–61
Printed in Belgium – all rights reserved

INTERNATIONAL JOURNAL OF
PEPTIDE & PROTEIN RESEARCH
ISSN 0367-8377

Sensitive fluorometric assay for the activity of chymosin

HIROO YONEZAWA [1], TETSUYA UCHIKOBA [1], MAKOTO KANEDA [1] and NOBUO IZUMIYA [2]

[1] *Department of Chemistry, Faculty of Science, Kagoshima University, Kagoshima, Japan and* [2] *Laboratory of Biochemistry, Faculty of Engineering, Kyushu Sangyo University, Fukuoka, Japan*

Received 13 April, revised 15 May, accepted for publication 28 July 1995

Fluorogenic substrates for chymosin [Dns-Leu-Ser-Phe-Trp-Ala-Leu-OCH_2Py (I), Dns-Leu-Ser-Phe-Met-Trp-Leu-OCH_2Py (II), Dns-Leu-Ser-Leu-Trp-Ala-Leu-OCH_2Py (III), Dns-Leu-Ala-Phe-Trp-Ala-Leu-OCH_2Py (IV), Dns-Leu-Ser-Phe-Leu-Ala-Leu-OCH_2Py (V) and Dns-Leu-Ser-Phe-Phe-Ala-Leu-OCH_2Py (VI)] were synthesized by a solution method. The obtained substrates I–VI were cleaved specifically (between the Phe and Trp residues for substrates I and IV, the Phe and Met residues for substrate II, the Leu and Trp residues for substrate III, the Phe and Leu residues for substrate V, and the Phe and Phe residues for substrate VI) by chymosin. The fluorescence of substrates I–IV (345 nm) increased with their hydrolysis, and hydrolysis rates were obtained by measuring the increase in fluorescence. The minimum detectable chymosin concentrations for substrates I and IV were about 1 nM; those for substrates II and III were about 4 and 2 nM. This assay method is very sensitive, and it is possible to determine the chymosin activity rapidly and easily. Substrates I and IV–VI were hydrolyzed by chymosin two times faster than substrates II and III. The effect of the amino-acid residues of the substrates on the hydrolysis rate is discussed.

Key words: chymosin; dansyl peptide; fluorogenic substrate; protease activity

Chymosin is a milk-clotting enzyme from the fourth stomach of the calf; the enzyme specifically cleaves the Phe^{105}–Met^{106} bond of κ-casein (1). Chymosin activity determinations rely on its clotting activity on milk. However, these determinations have two disadvantages, viz. they are indirect methods and they depend on a clotting process controlled by a series of complex factors. Moreover, some synthetic substrates have been used (2, 3). However, a good chromogenic substrate for chymosin has not been used. Raymond *et al.* synthesized chromogenic substrates containing *p*-nitrophenylalanine at the P_1 position, but the sensitivity of the substrates for chymosin action was low (4, 5). On the other hand, fluorogenic substrates have been used for highly sensitive assay methods for many protease activities (6–8). In our previous paper we reported the synthesis of intramolecularly quenching substrates for pepsin and the measurement of pepsin activity (9, 10). In this method it is possible to determine the pepsin activity rapidly and simply.

It was stated by Raymond *et al.* that Leu-Ser-Phe(NO_2)-Nle-Ala-Leu-OMe was cleaved rapidly by chymosin (4); accordingly, we synthesized some analogs which contain a dansyl group at the *N*-terminus of the peptide and a Trp residue at position P_1' or P_2'. These substrates contain many hydrophobic amino-acid residues; therefore, the *N*-protected peptides may be not sufficiently soluble in acidic buffers. Then, in order to increase the solubility of the substrates in the acidic buffers, a cationic group (4-pyridylmethyl ester) was introduced at the *C*-terminal carboxyl group of the substrates (11).

EXPERIMENTAL PROCEDURES

Synthesis of substrates

Boc-Ser(Bzl)-Phe-OBzl (I). To a chilled solution of Boc-Ser(Bzl) (1.48 g, 5 mmol), Phe-OBzl·TosOH (2.13 g, 5 mmol) and Et_3N (0.7 mL, 5 mmol) in CH_2Cl_2 (50 mL), was added DCC (1.03 g, 5 mmol). The reaction mixture was stirred for 1 h at 0 °C and overnight at room tem-

Abbreviations used: Dns, dansyl; Py, 4-pyridyl; Nle, L-norleucine; DCC, dicyclohexylcarbodiimide; EEDQ, 1-ethoxycarbonyl-1,2-ethoxy-2-hydroquinoline; AcOH, acetic acid; DMF, dimethylformamide; TCA, trichloroacetic acid.

perature, then evaporated *in vacuo*, and ethyl acetate was added to the residue. After dicyclohexylurea was filtered off, the filtrate was washed successively with 10% citric acid, 4% sodium bicarbonate solution and water. The organic layer was dried over anhydrous sodium sulfate. It was evaporated *in vacuo* and the crystals were collected by filtration with the aid of petroleum ether. The product was recrystallized from ethyl acetate–petroleum ether; yield 2.36 g (89%); m.p. 83–84 °C; $[\alpha]_D^{20}$ −4.5 ° (*c* 1, ethanol). Found: C, 70.06; H, 6.73; N, 5.13%. Calcd. for $C_{31}H_{36}O_6N_2$: C, 69.90; H, 6.81; N, 5.26%.

Z-Leu-Ser(Bzl)-Phe-OBzl (II). Compound **I** (1.06 g, 2 mmol) was treated with 10 mL of 1 M HCl/AcOH for 1 h and evaporated *in vacuo*, and the crystals were collected by filtration with the aid of petroleum ether. The resulting HCl·Ser(Bzl)-Phe-OBzl and Z-Leu·DCHA (1.12 g, 2.5 mmol) were dissolved in CH_2Cl_2 (20 mL) and coupled in the same manner using DCC as described above. The product was recrystallized from DMF–ether; yield 1.25 g (92%); m.p. 162–163 °C; $[\alpha]_D^{20}$ −5.8 ° (*c* 1, DMF). Found: C, 70.44; H, 6.89; N, 6.01%. Calcd. for $C_{40}H_{45}O_7N_3$: C, 70.67; H, 6.67; N, 6.18%.

Leu-Ser-Phe (III). Compound **II** (1.22 g, 1.8 mmol) was suspended in a mixture of methanol–AcOH–water (8:2:1, 30 mL) and hydrogenated in the presence of palladium black. After 24 h the catalyst was filtered off, the filtrate was evaporated *in vacuo* and the crystals were collected by filtration with the aid of acetone. The product was recrystallized from 50% AcOH–acetone; yield 0.40 g (58%); m.p. 244–246 °C; $[\alpha]_D^{20}$ 31.6 ° (*c* 1, 50% AcOH). Found: C, 56.15; H, 7.86; N, 10.72%. Calcd. for $C_{18}H_{27}O_5N_3 \cdot H_2O$; C, 56.38; H, 7.62; N, 10.96%.

Dns-Leu-Ser-Phe (IV). Compound **III** (350 mg, 0.91 mmol) was dansylated with dansyl chloride using a procedure similar to that of Gray (12). The product was recrystallized from ethanol–ether; yield 380 mg (68%); m.p. 200–202 °C; $[\alpha]_D^{20}$ −12.3 ° (*c* 1, DMF). Found: C, 58.65; H, 6.48; N, 8.86%. Calcd. for $C_{30}H_{38}O_7N_4S_1 \cdot H_2O$: C, 58.42; H, 6.54; N, 9.09%.

Dns-Leu-Ala-Phe (V). Leu-Ala-Phe (350 mg, 1 mmol) was dansylated in the same manner as described above. The product was recrystallized from ethanol–ether; yield 490 mg (84%); m.p. 196–198 °C; $[\alpha]_D^{20}$ −2.0 ° (*c* 1, DMF). Found: C, 61.67; H, 6.36; N, 9.79%. Calcd. for $C_{30}H_{38}O_6N_4S_1$: C, 61.83; H, 6.57; N, 9.62%.

Dns-Leu-Ser-Leu (VI). This was prepared from Leu-Ser-Leu (350 mg, 1 mmol) in the same manner as described above. The product was recrystallized from ethanol–ether; yield 360 mg (62%); m.p. 130–132 °C; $[\alpha]_D^{20}$ −29.6 ° (*c* 1, DMF). Found: C, 55.82; H, 7.41; N, 9.78%. Calcd. for $C_{27}H_{40}O_7N_4S_1 \cdot H_2O$: C, 55.65; H, 7.27; N, 9.62%.

Boc-Ala-Leu-OCH₂Py (VII). Boc-Leu-OCH_2Py (9) (2.00 g, 6.2 mmol) was treated with 1 M HCl/AcOH for 1 h at room temperature, and the resulting HCl·Leu-OCH_2Py was converted into free base as described previously (9). The obtained oily Leu-OCH_2Py and Boc-Ala (1.13 g, 6 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using 1-ethoxycarbonyl-1,2-ethoxy-2-hydroquinoline (EEDQ) (1.48 g, 6 mmol); yield of oil 2.03 g (86%).

Boc-Trp-Ala-Leu-OCH₂Py (VIII). Compound **VII** (2.00 g, 5.1 mmol) was treated with 1 M HCl/AcOH, and the resulting HCl·Ala-Leu-OCH_2Py was converted into free base as described above. The obtained oily Ala-Leu-OCH_2Py and Boc-Trp (1.52 g, 5 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using EEDQ. The product was recrystallized from ethyl acetate–petroleum ether; yield 2.39 g (82%); m.p. 95–97 °C; $[\alpha]_D^{20}$ −29.0 ° (*c* 1, ethanol). Found: C, 64.05; H, 7.38; N, 11.89%. Calcd. for $C_{31}H_{41}O_6N_5$: C, 64.23; H, 7.13; N, 12.08%.

Boc-Trp-Leu-OCH₂Py (IX). Leu-OCH_2Py (6.2 mmol) and Boc-Trp (1.82 g, 6 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using EEDQ; yield of oil 2.44 g (80%).

Boc-Met-Trp-Leu-OCH₂Py (X). Compound **IX** (2.40 g, 4.7 mmol) was treated with 1 M HCl/AcOH containing 2% 2-mercaptoethanol, and the resulting HCl·Trp-Leu-OCH_2Py was converted into free base as described above. The obtained oily Trp-Leu-OCH_2Py and Boc-Met (1.17 g, 4.7 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using EEDQ. The product was recrystallized from ethyl acetate–petroleum ether; yield 2.53 g (84%); m.p. 117–119 °C; $[\alpha]_D^{20}$ −31.1 ° (*c* 1, ethanol). Found: C, 61.68; H, 7.04; N, 10.71%. Calcd. for $C_{33}H_{45}O_6N_5S_1$: C, 61.95; H, 7.09; N, 10.94%.

Boc-Leu-Ala-Leu-OCH₂Py (XI). Ala-Leu-OCH_2Py (2.7 mmol) and Boc-Leu (630 mg, 2.7 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using EEDQ; yield of oil 1.13 g (83%).

Boc-Phe-Ala-Leu-OCH₂Py (XII). Ala-Leu-OCH_2Py (2.7 mmol) and Boc-Phe (720 mg, 2.7 mmol) were dissolved in CH_2Cl_2 and coupled in the same manner using EEDQ. The product was recrystallized from ethyl acetate–petroleum ether; yield 1.0 g (68%); m.p. 160–161 °C; $[\alpha]_D^{20}$ −31.9 ° (*c* 1, ethanol). Found: C, 64.21; H, 7.68; N, 10.15%. Calcd. for $C_{29}H_{40}O_6N_4$: C, 64.42; H, 7.46; N, 10.36%.

Dns-Leu-Ser-Phe-Trp-Ala-Leu-OCH₂Py (XIII). Compound **VIII** (174 mg, 0.3 mmol) was treated with 1 M

HCl/AcOH containing 2% 2-mercaptoethanol, and the resulting HCl·Trp-Ala-Leu-OCH_2Py was converted into free base as described above. The obtained oily Trp-Ala-Leu-OCH_2Py and Dns-Leu-Ser-Phe (170 mg, 0.28 mmol) were dissolved in CH_2Cl_2 (10 mL) and DMF (5 mL), and EEDQ (70 mg, 0.28 mmol) was added to the solution. The reaction mixture was stirred for 1 h at 0 °C and overnight at room temperature, evaporated *in vacuo* to remove CH_2Cl_2, and 4% sodium bicarbonate solution was added to the residue. The resulting precipitate was collected by filtration, washed with 10% citric acid and water and dried. The product was recrystallized from DMF–ether; yield 260 mg (86%); m.p. 235–237 °C; $[\alpha]_D^{20}$ −21.4° (*c* 1, DMF). Found: C, 62.51; H, 6.85; N, 11.47%. Calcd. for $C_{56}H_{69}O_{10}N_9S_1 \cdot H_2O$: C, 62.37; H, 6.64; N, 11.69%.

Dns-Leu-Ser-Phe-Met-Trp-Leu-OCH_2Py (XIV). This was prepared from Met-Trp-Leu-OCH_2Py (0.31 mmol) and Dns-Leu-Ser-Phe (180 mg, 0.3 mmol) in the same manner as described above. The product was recrystallized from DMF–ether; yield 210 mg (62%); m.p. 226–228 °C; $[\alpha]_D^{20}$ −99.1° (*c* 1, DMF). Found: C, 61.03; H, 6.89; N, 10.82%. Calcd. for $C_{58}H_{73}O_{10}N_9S_2 \cdot H_2O$: C, 61.19; H, 6.64; N, 11.07%.

Dns-Leu-Ser-Leu-Trp-Ala-Leu-OCH_2Py (XV). This was prepared from Trp-Ala-Leu-OCH_2Py (0.4 mmol) and Dns-Leu-Ser-Leu (170 mg, 0.3 mmol) in the same manner as described above. The product was recrystallized from DMF–ether; yield 120 mg (38%); m.p. 242–244 °C; $[\alpha]_D^{20}$ −33.7° (*c* 1, DMF). Found: C, 60.88; H, 7.31; N, 11.86%. Calcd. for $C_{53}H_{71}O_{10}N_9S_1 \cdot H_2O$: C, 60.96; H, 7.05; N, 12.07%.

Dns-Leu-Ala-Phe-Trp-Ala-Leu-OCH_2Py (XVI). This was prepared from Trp-Ala-Leu-OCH_2Py (0.4 mmol) and Dns-Leu-Ala-Phe (175 mg, 0.3 mmol) in the same manner as described above. The product was recrystallized from DMF–ethyl acetate; yield 220 mg (69%); m.p. above 250 °C; $[\alpha]_D^{20}$ −18.7° (*c* 1, DMF). Found: C, 63.04; H, 6.95; N, 11.62%. Calcd. for $C_{56}H_{69}O_9N_9S_1 \cdot H_2O$: C, 63.31; H, 6.74; N, 11.87%.

Dns-Leu-Ser-Phe-Leu-Ala-Leu-OCH_2Py (XVII). This was prepared from Leu-Ala-Leu-OCH_2Py (0.4 mmol) and Dns-Leu-Ser-Phe (180 mg, 0.3 mmol) in the same manner as described above. The product was recrystallized from DMF–ether; yield 255 mg (85%); m.p. 242–244 °C; $[\alpha]_D^{20}$ −28.4° (*c* 1, DMF). Found: C, 60.77; H, 7.01; N, 11.34%. Calcd. for $C_{51}H_{70}O_{10}N_8S_1 \cdot H_2O$: C, 60.93; H, 7.22; N, 11.15%.

Dns-Leu-Ser-Phe-Phe-Ala-Leu-OCH_2Py (XVIII). This was prepared from Phe-Ala-Leu-OCH_2Py (0.4 mmol) and Dns-Leu-Ser-Phe (180 mg, 0.3 mmol) in the same manner as described above. The product was recrystallized from DMF–ether; yield 150 mg (48%); m.p. 247–249 °C; $[\alpha]_D^{20}$ −24.9° (*c* 1, DMF). Found: C, 62.57; H, 6.83; N, 10.51%. Calcd. for $C_{54}H_{68}O_{10}N_8S_1 \cdot H_2O$: C, 62.41; H, 6.79; N, 10.78%.

Materials and enzymic studies

Chymosin (SIGMA Chemical Co., USA) was dissolved in H_2O. The substrates were dissolved in 0.1 M citrate buffer containing 10% of DMF. Fluorescence measurements were performed with a Shimadzu spectrofluorometer RF-5000.

Analysis of reaction products. Chymosin solution (10 μL, 10 μM) was added to 1 mL of the substrate solutions (50 μM, pH 2.5) and incubated at 37 °C. After incubation for 24 h, 0.1 mL of 3 M TCA solution was added to the reaction mixtures, the resulting precipitates were removed by filtration, and 100 μL of the filtrates were subjected to HPLC assay.

Measurement of the fluorescence emission spectrum. Chymosin solution (50 μL, 1 μM) was added to 5 mL of solutions of substrates I–IV (50 μM, pH 2.5). After

TABLE 4

Kinetic constants of the substrates

P_3 P_2 P_1 P_1' P_2' P_3'		Optimum pH	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Dns-Leu-Ser-Phe-Trp-Ala-Leu-OCH_2Py (I)	C[a]	2.5	30.2 ± 2.8	4.27 ± 0.40	0.14
	P[a]	3.5	28.2 ± 3.2	3.85 ± 0.19	0.13
Dns-Leu-Ser-Phe-Met-Trp-Leu-OCH_2Py (II)	C	2.0	14.6 ± 1.2	1.05 ± 0.07	0.068
	P	2.0	16.2 ± 1.9	0.96 ± 0.08	0.059
Dns-Leu-Ser-Leu-Trp-Ala-Leu-OCH_2Py (III)	C	2.5	16.0 ± 1.8	1.10 ± 0.06	0.069
	P	3.5	40.0 ± 2.7	6.67 ± 0.43	0.17
Dns-Leu-Ala-Phe-Trp-Ala-Leu-OCH_2Py (IV)	C	2.0	30.1 ± 3.2	5.56 ± 0.35	0.18
	P	2.0	20.0 ± 1.8	2.38 ± 0.14	0.12
Dns-Leu-Ser-Phe-Leu-Ala-Leu-OCH_2Py (V)	C	2.5	35.2 ± 8.7	3.86 ± 0.68	0.11
Dns-Leu-Ser-Phe-Phe-Ala-Leu-OCH_2Py (VI)	C	2.5	30.4 ± 7.2	4.16 ± 0.59	0.14

[a] C, chymosin; P, pepsin.

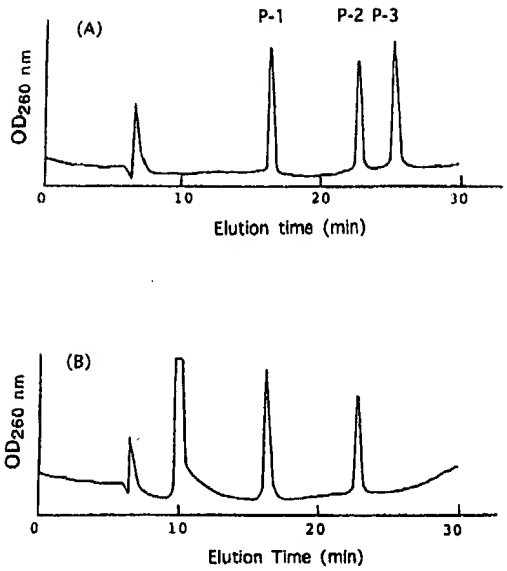


FIGURE 1

HPLC analysis of substrate I and reaction mixture. (A) Mixture of substrate I, Dns-Leu-Ser-Phe and Trp-Ala-Leu-OCH_2Py. P-1, Trp-Ala-Leu-OCH_2Py; P-2, Dns-Leu-Ser-Phe; P-3, substrate I. (B) Reaction mixture of substrate I and chymosin. The reaction mixture (100 μL) was applied to a C_{18} reversed-phase column (Dynamax-60A, 4.6 x 250 mm, Rainin Instruments) and eluted with a linear gradient of 20–80% 2-propanol/acetonitrile (7/3, v/v) containing 0.1% trifluoroacetic acid in 0.1% trifluoroacetic acid for 30 min at a flow rate of 0.5 mL min^{-1}.

these solutions had been incubated for 10 s and 24 h, the fluorescence emission spectrum was measured (excitation at 290 nm).

pH–Activity curve. Substrates I–IV were dissolved in 5 mL of the buffers in the pH range 1.5–3.5 (50 μM), and chymosin solutions (20 μL of 1 μM solution for substrates I and IV, 50 μL for substrates II and III) were added to the substrate solutions. After incubation for 1 min at 37 °C, the increase in fluorescence (345 nm, excitation at 290 nm) was recorded during the following 6 min of incubation. On the other hand, the optimum pH values of substrates V and VI were determined by HPLC assay.

Measurement of kinetic parameters. The kinetic parameters of substrates I–IV were determined by fluorometric assay at their optimum pH as described above, and those of substrates V and VI were determined by HPLC assay.

Relationship of the increase of fluorescence to enzyme concentration. Various amounts of the chymosin solutions were added to 5 mL solutions of substrates I–IV (25 μM, pH 2.0 or 2.5) and the increase in fluorescence was measured as described above.

RESULTS AND DISCUSSION

The substrates were obtained with the coupling of dansyl tripeptides and tripeptide 4-pyridylmethyl esters. The substrates obtained were incubated with chymosin for 24 h and the reaction mixtures were applied to HPLC. As shown in Figs. 1A and 1B, substrate I disappeared completely, and two products appeared (Dns-Leu-Ser-Phe and Trp-Ala-Leu-OCH_2Py). Similarly, other substrates were cleaved completely (between residues Phe and Met for substrate II; residues Leu and Trp for substrate III; residues Phe and Trp for substrate IV, residues Phe and Leu for substrate V, and residues Phe and Phe for substrate VI) by chymosin. This indicated that a detectable racemization reaction did not occur. As shown in Fig. 2, the fluorescence at 345 nm increased upon hydrolysis of the substrates.

The increased amount of organic solvent seemed to inhibit chymosin activity. In a preliminary experiment, chymosin activity for substrate I assayed in the presence of 10% DMF at pH 2.5 was found to be 70% of that in the presence of 5% DMF. Since the solubilities of these substrates were poor, particularly at high pH values, their activities were assayed in the presence of 10% DMF. Under these conditions the chymosin activity may be about 50–60% of that in buffers which are free from organic solvent.

Comparative measurements of the effect of pH on chymosin action were made with the substrates. As shown in Fig. 3, the optimum pH for substrates II and IV was 2.0, and for I, III, V and VI it was 2.5. Thus pH 2.0 or 2.5 buffers were used as solvents for the substrates in the following experiments.

When chymosin solutions were added to solutions of substrates I–IV the fluorescence at 345 nm increased linearly in proportion to the concentration of added chymosin solution (Fig. 4). The minimum detectable chymosin concentrations for substrates I and IV were 1.0 nM (the curve for substrate IV was the same as that for substrate I), and those of substrates II and III were 4 and 2 nM, respectively.

Substrate I was hydrolyzed two times faster than substrate II (Table 1). In order to elucidate the reason for the slow rate of hydrolysis of substrate II, Dns-Leu-Ser-Phe-Leu-Ala-Leu-OCH_2Py (substrate V) and Dns-Leu-Ser-Phe-Phe-Ala-Leu-OCH_2Py (substrate VI) were synthesized. As shown in Table 1, substrates I, V and VI were hydrolyzed at nearly the same rate. This indicates that replacing the Met residue with a Phe residue at the P_1 position does not have significant influence on the hydrolysis rate. On the other hand, substrate II was hydrolyzed two times slower than substrate V. This indicates a dislike for a large aromatic side chain at position P_2.

The effect of position P_1 on the hydrolysis rate was then examined. Substrate I (with Phe at P_1) was hydrolized faster than substrate III (with Leu at P_1). This

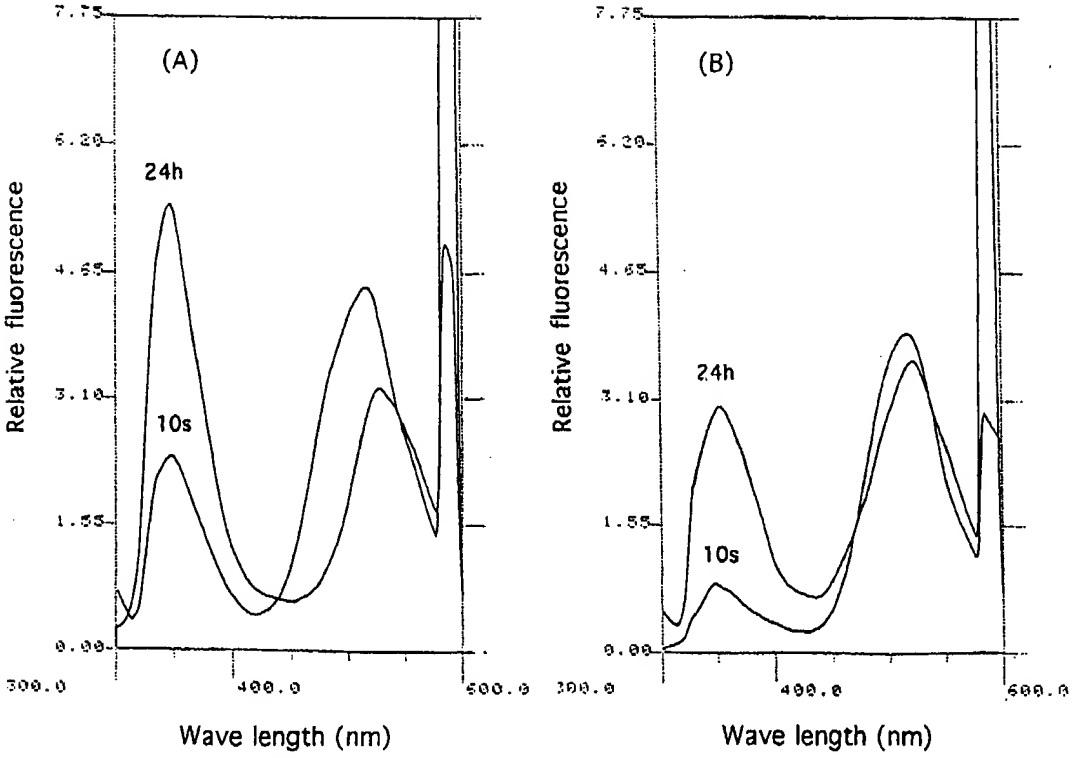


FIGURE 2

Fluorescence emission spectra of substrates I and II before and after hydrolysis. (A) Substrate I and chymosin solution were mixed, and after 10 s and 24 h, the fluorescence emission spectrum was measured. (B) The fluorescence emission spectrum of substrate II was measured in the same manner.

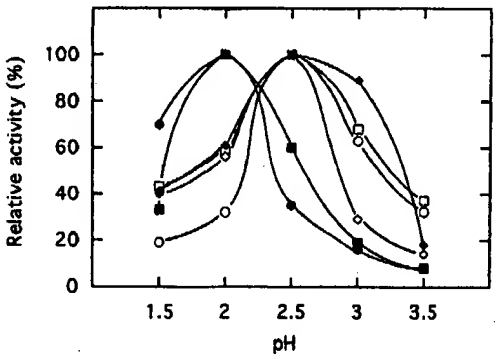


FIGURE 3

pH Dependence of hydrolysis of the substrates. The substrates were dissolved in buffers of various pH values, and chymosin solutions were added. After incubation for 1 min, the increase in fluorescence was measured during the following 6 min. (○) Substrate I; (●) substrate II; (□) substrate III; (■) substrate IV; (◆) substrate V; (◇) substrate VI.

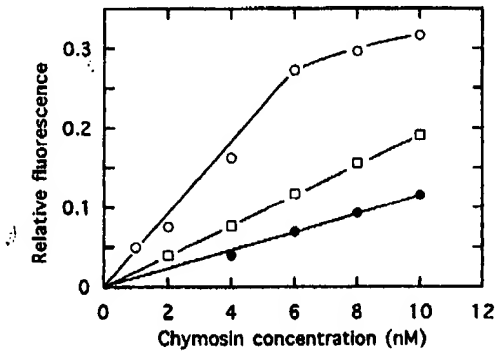


FIGURE 4

Increase in fluorescence vs. chymosin concentration. Various amounts of chymosin solutions were added to the substrate solutions and the increase in fluorescence was measured. (○) substrate I and IV; (●) substrate II; (□) substrate III.

shows that chymosin exhibits a preference for the presence of an aromatic side chain at P_1.

Moreover, the effect of position P_2 on the hydrolysis rate was examined.

Substrates I and IV were hydrolyzed at nearly the same rate. The amino-acid residue at the P_2 position

did not have a significant influence on the hydrolysis rate.

These substrates were hydrolyzed by pepsin at nearly the same rates for chymosin (Table 1). Moreover, substrates **I–IV** were hydrolyzed by bovine spleen cathepsin D, but the hydrolysis rates were slow (the minimum detectable concentration was 20–50 nM). This assay method for chymosin activity is very sensitive and convenient; therefore, this method will be useful for research into the action of chymosin.

REFERENCES

1. Jolles, J., Alais, C. & Jolles, P. (1968) *Biochem. Biophys. Acta* **168**, 591–593
2. Hill, R.D. (1969) *J. Dairy Res.* **36**, 409–415
3. Polzhofer. (1972) *Tetrahedron* **28**, 855–865
4. Raymond, M.N., Bricas, E., Salesse, R., Garnier, J., Garnot, P. & Ribadeau Dumas, B. (1973) *J. Dairy Sci.* **56**, 419–422
5. Raymond, M.N. & Bricas, E. (1979) *J. Dairy Sci.* **62**, 1719–1725
6. Greenberg, L.J. (1962) *Biochem. Biophys. Res. Commun.* **9**, 430–435
7. Nagatsu, I., Nagatsu, T., Yamamoto, T., Glenner, G.G. & Mehi, J.W. (1970) *Biochem. Biophys. Acta* **198**, 255–270
8. Matayoshi, E.D., Wang, G.T., Krafft, G.A. & Erickson, J. (1990) *Science* **247**, 954–958
9. Yonezawa, H. & Izumiya, N. (1991) *Bull. Chem. Soc. Jpn.* **64**, 1407–1409
10. Yonezawa, H., Yamada, K., Uchikoba, T. & Kaneda, M. (1993) *Bull. Chem. Soc. Jpn.* **66**, 2653–2658
11. Sachdev, G.P. & Fruton, J.S. (1969) *Biochemistry* 8,4231–4238
12. Gray, W.R. (1967) *Meth. Enzymol.* **11**, 139–151

Address:

Dr. *Hiroo Yonezawa*
Department of Chemistry
Faculty of Science
Kagoshima University
1-Korimoto
Kagoshima 890
Japan

ELSEVIER

Journal of Biotechnology 45 (1996) 235–241

JOURNAL OF Biotechnology

Investigating the use of the chymosin-sensitive sequence of κ-casein as a cleavable linker site in fusion proteins [1]

Marie K. Walsh [2], Harold E. Swaisgood *

Department of Food Science, North Carolina State University, Raleigh, NC 27695-7624, USA

Received 7 August 1995; revised 9 November 1995; accepted 9 November 1995

Abstract

The chymosin-sensitive sequence of bovine k-casein A (κ-CN A) was investigated as a cleavable linker site between the two domains of a streptavidin-chloramphenicol acetyltransferase fusion protein. Two DNA sequences were synthesized which encode the amino acids from 101 to 107 and from 97 to 113 of bovine κ-CN A. These sequences were separately cloned in-frame to a streptavidin expression vector used for fusion protein construction. The gene for chloramphenicol acetyltransferase (CAT) was then cloned in-frame to a streptavidin-chymosin-sensitive linker vector forming plasmids pStCL1CAT and pStCL2CAT. The fusion protein was expressed in *Escherichia coli* and SDS-PAGE and Western blot analysis of chymosin-treated cell lysates showed a pH-dependent cleavage of the fusion proteins. Fusion proteins were also bioselectively immobilized onto biotinylated controlled-pore glass beads and treated with chymosin. CAT was specifically released by chymosin treatment and was identified by SDS-PAGE.

Keywords: Expression vector; Cleavable linker; Fusion protein; Affinity purification; Chymosin

1. Introduction

Recombinant DNA techniques are frequently used to construct bifunctional proteins. These fusion proteins are often designed to contain an affinity tail fused to the protein of interest to facilitate isolation of the recombinant protein. Our laboratory is currently using streptavidin as an affinity domain to bioselectively immobilize recombinant proteins from crude cell lysates. At times, a soluble gene product is needed, therefore site-specific cleavage of the fusion protein between the affinity tail and the protein of interest is necessary.

Several recognition sequences have previously been employed, including that for collagenase, factor Xa, thrombin and enterokinase (reviewed in Uhlen and Moks, 1990). A specific protease must be used to limit proteolysis to the linker site only, and this site must be designed so that the recognition sequence is accessible to the protease.

Chymosin (EC 3.4.23.04) is an aspartyl protease used for milk clotting (Foltmann, 1981; Fox, 1993). The primary action of chymosin in the coagulation

* Corresponding author.

[1] Paper No. FSR-95-28 of the Journal Series of the Department of Food Science, North Carolina State University, Raleigh, NC 27695-7624. The use of trade names in this publication does not imply endorsement by the North Carolina Agricultural Research Service of products named, nor criticism of similar ones not mentioned.

[2] Current address: Utah State University, Department of Nutrition and Food Sciences, Logan, UT 84322-8700, USA.

SSDI 0168-1656(95)00178-6

of milk is the specific hydrolysis of bovine κ-CN at the Phe_{105}–Met_{106} bond, in fact, κ-casein was the only component of whole casein attacked by chymosin within 50 min at pH 6.8 (El-Negoumy, 1968). The optimum pH for general proteolysis by chymosin is between pH 3 and 4.0 (Foltmann, 1981; Keil, 1992); however, cleavage at pH 6.8 is specific for the Phe–Met bond in κ-CN. Chymosin is the most specific of the aspartyl proteases, with the specific hydrolysis of κ-casein being dependent upon the primary, and probably, the secondary and tertiary structure of an extended region around the susceptible bond (Dalgleish, 1982; Fox, 1993).

Chymosin has been cloned, and recombinant enzyme costs about half as much as the calf enzyme (De Palma, 1994). We have investigated the use of the chymosin-sensitive sequence in bovine κ-CN as a cleavable linker site for production of soluble recombinant protein.

2. Materials and methods

2.1. Materials

Plasmid pStp4, containing a tac promoter followed by the coding region for streptavidin minus the stop codon (Walsh and Swaisgood, 1994), was used as the parent vector. Restriction enzymes were from New England Biolabs (Beverly, MA). Isolation of plasmid DNA, restriction enzyme digestions and transformations of competent cells were according to manufacturers' recommendations or Sambrook et al. (1989). *E. coli* NM522 cells were from Invitrogen (San Diego, CA) and plasmid pCAT-Basic vector containing the coding region for CAT was obtained from Promega (Madison, WI). Oligonucleotides were synthesized at the Molecular Biology Center (North Carolina State University, Raleigh, NC). Other chemicals and supplies were from Sigma (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA) unless otherwise noted.

2.2. Vector constructions

Oligonucleotides containing varying lengths of the coding region for the chymosin-sensitive sequence of bovine κ-CN a were mixed together in a 1:1 ratio, heated to 100°C for 5 min and slowly cooled to room temperature. Hybridized oligonucleotides were cloned into plasmid pStp4 using *Sal*I and *Bam*HI restriction enzyme sites. These oligonucleotides also contain a unique *Eco*RI restriction enzyme site. The gene for CAT was excised from pCAT-Basic and cloned in-frame to the chymosin linker vectors using *Hpa*I and *Xba*I, thus creating plasmids pStCL1CAT and pStCL2CAT. Vectors were transformed into *E. coli* NM522 cells.

2.3. Protein expression and bioreactor loading

E. coli NM522 cells containing either expression construct were grown to an optical density of 0.6 at 600 nm in NZYM (Sambrook et al., 1989) supplemented with 50 mg ml^{-1} ampicillin (Sigma, St. Louis, MO) at 37°C. IPTG (Promega, Madison, WI) was added to give a final concentration of 1 mM to induce expression. Cells were further incubated for 4 h for protein production before harvesting by centrifugation at $4000 \times g$ and 4°C for 10 min. Cells were resuspended in 50 mM Tris buffer, pH 7.8, supplemented with 3 mM EDTA and ruptured by sonication. The lysates were centrifuged at $8000 \times g$ and 4°C to precipitate insoluble debris. Crude cell lysates were assayed for CAT activity using a spectrophotometric method (Shaw, 1975).

Aminopropyl controlled-pore glass (CPG) beads were prepared and biotinylated with NHS-LC Biotin (Pierce, Rockford, IL) as previously described (Walsh and Swaisgood, 1994). Crude cell lysates from induced cells transformed with pStCL1CAT or pStCL2CAT were each passed through 0.5 ml of the biotinylated beads to separately immobilize the two streptavidin-CAT fusion proteins. Approx. 4 l of lysed cells were used per 0.5 ml biotinylated beads. Beads were washed with 2 M NaCl after exposure to each liter of lysate, then washed with 4 M urea containing 2 M NaCl after adsorption from all 4 l to remove nonspecifically adsorbed proteins. Beads were treated with 10 U of chymosin (Sigma, St. Louis, MO) at pH 4.0, 5.3 and 6.8 for 4 and 12 h. Beads were then washed in 2 M NaCl and the eluant was assayed for CAT using a CAT-ELISA (Boehringer Mannheim, Indianapolis, IN) and assayed for protein composition by SDS-PAGE.

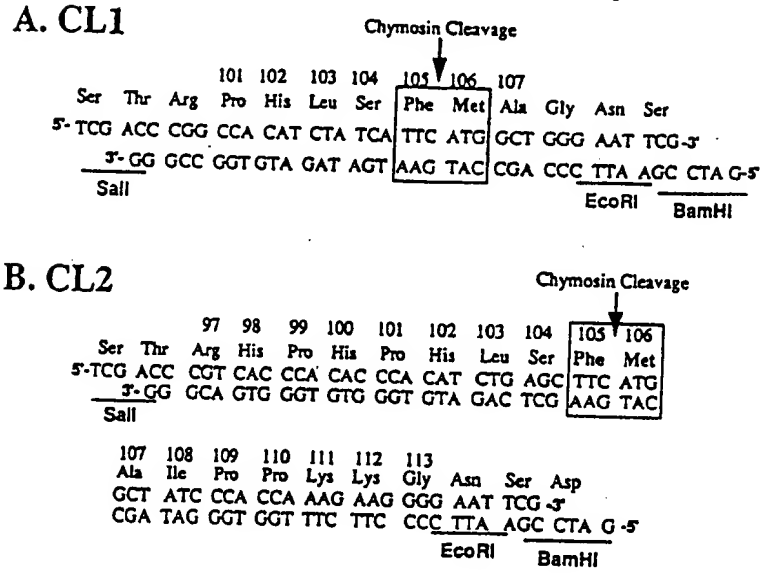


Fig. 1. Nucleotide sequences of chymosin-sensitive linker oligonucleotides. Both sequences contain unique *SalI*, *EcoRI* and *BamHI* restriction enzyme sites and the Phe–Met bond recognized by chymosin. (A) Chymosin linker 1 (CL1) contains amino acid residues 101 (proline) to 107 (alanine) from κ-casein A. (B) Chymosin linker 2 (CL2) contains amino acid residues 97 (arginine) to 113 (glycine) from κ-casein A.

2.4. Soluble chymosin assay

Soluble protein from induced lysed cells was heat treated at 75°C, centrifuged and the pellet discarded. Remaining protein was assayed for CAT activity using the spectrophotometric method (Shaw, 1975) and total protein was determined spectrophotometrically using the Bio-Rad protein assay (Bio-Rad, Hercules, CA) with bovine serum albumin as the standard. Protein solutions were adjusted to 1 mg/ml in citrate buffer pH 4.0, 5.3 or 6.8. Chymosin, 6 U, was added and samples were incubated at 25°C for 12 h. Samples were analyzed by both SDS-PAGE with Coomassie blue staining and Western blot using

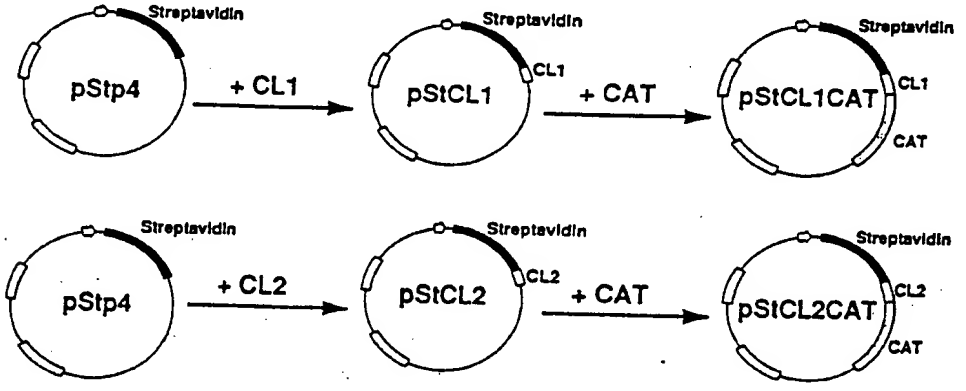


Fig. 2. Schematic drawing of streptavidin-chymosin-sensitive linker-CAT vectors. Chymosin-sensitive linkers were cloned in-frame to streptavidin in pStp4 using *SalI* and *BamHI* restriction enzyme sites. CAT was then cloned in frame to both chymosin-sensitive linker vectors using *HpaI* and *XhoI* restriction enzyme sites.

anti-streptavidin antibody (Walsh and Swaisgood, 1994).

3. Results and discussion

3.1. Linker design and vector construction

The chymosin-sensitive linkers shown in Fig. 1 contain the coding regions from amino acids 101 to 107 (CL1) or 97 to 113 (CL2) of bovine κ-CN A. These two sequences were used to determine the effects of linker length on chymosin recognition and subsequent cleavage. CL1 was tested to determine the minimum sequence needed for chymosin cleavage and CL2 was used to determine if an increase in linker length and/or linker secondary structure had effects on susceptibility to chymosin cleavage. Both sequences contain the Phe_{105}–Met_{106} bond which is recognized by chymosin. The linkers also contain unique restriction enzyme sites, *Sal*I and *Bam*H1 used for cloning, and an *Eco*RI site which was used

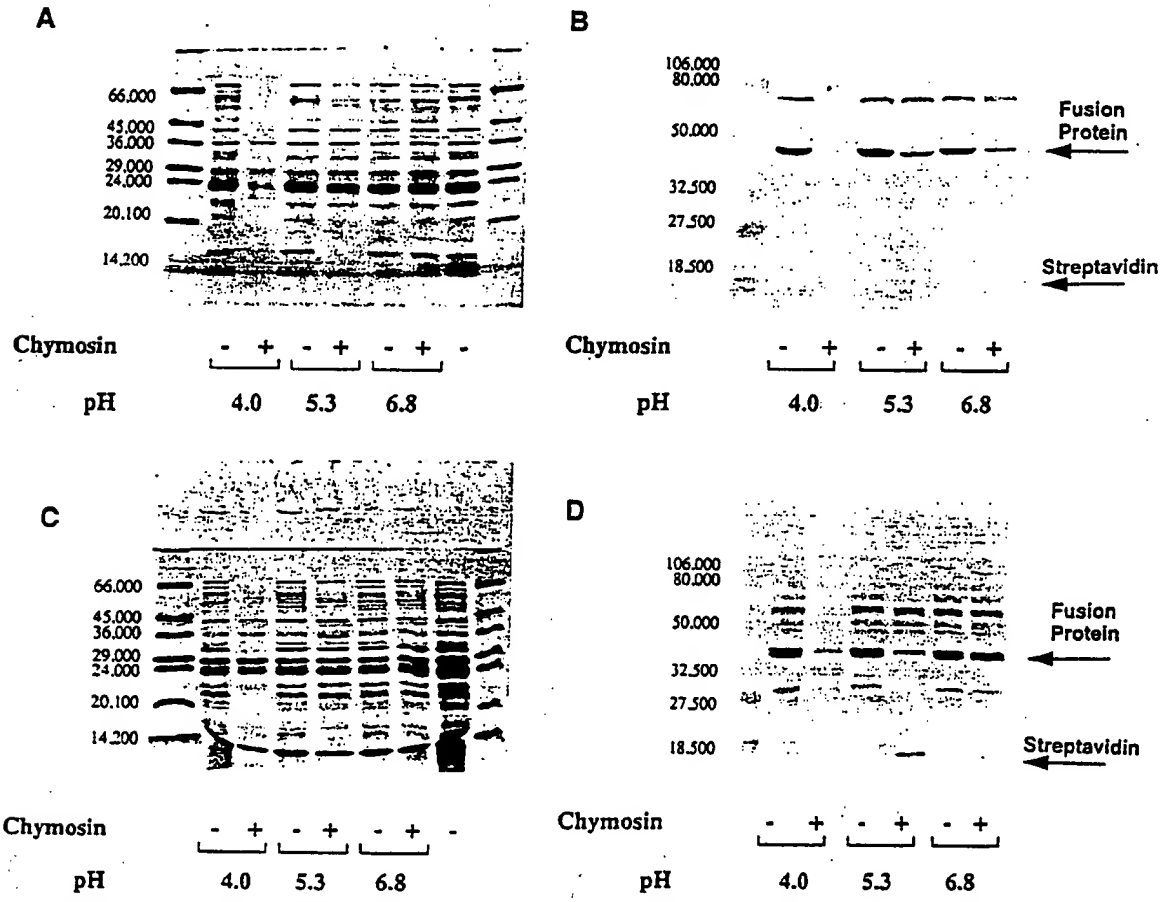


Fig. 3. SDS-PAGE and Western blot analysis of total protein from *E. coli* NM522 cells carrying either plasmid pStCL1CAT (A and B) or pStCL2CAT (C and D). (A and C) SDS-PAGE analysis of chymosin-treated cell lysates. Heat stable proteins were either treated with chymosin or left untreated at pH 4.0, 5.3, or 6.8. Lanes 1 and 9, molecular mass markers; lanes 2, 4, 6, untreated samples; lanes 3, 5, 7, chymosin-treated samples; lane 8, untreated sample, pH 7, kept at 0°C. (B and D) Western analysis of chymosin-treated heat stable proteins. The same samples run in A and C were run in the Western blots. Lane 1, molecular weight markers; lanes 2, 4, 6, untreated samples; lanes 3, 5, 7, chymosin-treated samples. Arrows indicate positions of the fusion proteins and released streptavidin.

to screen for positive colonies. Each vector contains 34 unique restriction enzyme sites following the chymosin-sensitive linker sequence to facilitate cloning a gene of interest in-frame to the streptavidin-CL sequences.

Fig. 2 shows the construction of streptavidin-CL1-CAT and streptavidin-CL2-CAT fusion protein vectors. Plasmid pStp4 contains the streptavidin gene, minus the stop codon, under the control of the tac promoter. The polylinker region contains 37 unique restriction enzyme sites in the polylinker following the coding region for streptavidin. The chymosin-sensitive linker oligonucleotides were separately cloned in-frame to the streptavidin gene, followed by cloning the CAT gene excised from pCAT-Basic. The resultant expression vectors are approx. 5.7 kb.

3.2. Soluble chymosin assay

Crude cell lysates from pStCL1CAT and pStCL2CAT transformants showed activities of 2.8 and 3.7 nmol chloramphenicol acetylated per min per mg protein. Western analysis of crude cell lysates treated with chymosin was done to qualitatively investigate the effects of both pH and linker length on chymosin cleavage. Analysis was performed on heat-treated crude cell lysates because both CAT and streptavidin are heat-stable proteins and the heat treatment removed contaminating proteins and inactivated native proteases. In samples treated at pH 4.0, there was a noticeable precipitate after 12 h. Because this pH is near the isoelectric point of both streptavidin and CAT, this precipitate may have been fusion protein. The isoelectric point of CAT is between 5.4 and 4.0 (Shaw, 1975), while that for streptavidin is between 5 and 6 (Green, 1990).

Western and SDS-PAGE analyses of heat-treated crude protein from lysed cells containing either pStCL1CAT or pStCL2CAT treated with chymosin at varying pH levels are shown in Fig. 3. The Western blots were treated with anti-streptavidin antibody. The calculated molecular mass of both fusion proteins are approx. 49 kDa. Chymosin treatment of the fusion protein should cleave the 49 kDa protein into two proteins of approx. 19 and 30 kDa. It is difficult to detect the appearance of the 19 and 30 kDa protein bands by SDS-PAGE analysis of heat-treated cell lysates, possibly because of overlapping bands. The absence of many Coomassie blue stained protein bands in the chymosin-treated samples at pH 4 suggests that more nonspecific proteolysis was occurring. General proteolytic activity of chymosin increases with a decrease in pH, with a maximum measured activity between pH 3 to 4 (Foltmann, 1981; Keil, 1992). Also, Western blot analysis of crude protein from pStCL1CAT and pStCL2CAT transformed cells detected very little fusion protein or streptavidin in samples treated with chymosin at pH 4.0.

Effects of linker length and pH on specific chymosin cleavage of the fusion proteins is shown in the Western analysis (Fig. 3). Analysis of fusion protein from pStCL1CAT transformed cells shows that very little cleavage occurred at either pH 5.3 or 6.8 as indicated by the lack of free streptavidin. (It should be noted that the intensities of bands in the Western blots are qualitative because of many variables that are difficult to control, such as the amount of protein in the band within the gel, transfer times, the amount of protein adsorbed to the membrane, and variability of photographic reproduction.) On the other hand, there is a dramatic reduction in the fusion protein band, with a subsequent increase in free streptavidin, upon chymosin treatment at pH 5.3 of the protein from pStCL2CAT transformed cells. With this fusion protein, specific cleavage occurred at both pH 6.8 and 5.3, with appreciably more occurring at pH 5.3. Both the decrease in the 49 kDa band and the increase in the low molecular mass band recognized by anti-streptavidin antibody indicated that CL2 was recognized and cleaved by chymosin more efficiently than CL1.

Specific hydrolysis of Phe_{105}–Met_{106} of κ-casein at pH 6.8 by chymosin is dependent upon the composition and sequence of amino acid residues in an extended region of the primary structure (Visser et al., 1976, 1980; Dalgleish, 1982). The accessibility and structure of this region are most likely important; for example, secondary and tertiary structure predictions suggest that this region of the protein protrudes from the surface of the molecule (Kumosinski et al., 1991). The effect of chymosin-sensitive linker length on the cleavage of fusion protein is consistent with the effect of peptide length on chymosin activities with model peptides. Comparison of the pseudo first-order rate constant, k_{cat}/K_m, for peptide

101–108 with that for peptide 98–112 indicated a 22-fold increase resulted when the larger peptide was substrate (Visser et al., 1976, 1980; Dalgleish, 1982). The pH optimum for cleavage of peptide 98–112 was 5.3–5.5 in agreement with the observed increase in cleavage of fusion protein at pH 5.3. In this study, the linker regions correspond to residues 101–107 and 97–113, with the longer linker resulting in appreciably more cleavage.

3.3. Chymosin-treated immobilized fusion proteins

Chymosin-treated fusion proteins were bioselectively immobilized onto biotinylated CPG beads and analyzed for CAT using a CAT-ELISA. The results are shown in Table 1. Immobilized streptavidin-CL1-CAT and streptavidin-CL2-CAT yielded CAT activities of 2205 and 2402 U/mg beads. Although the immobilized CAT activity was the same for both fusion proteins, a 12-h incubation with chymosin for StCL2CAT fusion protein yielded a much larger amount of released CAT. From Western analysis, treatment of CL2 at pH 5.3 appeared to result in more cleavage of the fusion protein than treatment at pH 6.8. Since pH 5.3 is near the isoelectric point of CAT, some of the released CAT may have aggregated and therefore was not efficiently detected by the CAT-ELISA or possibly aggregated CAT was retained within the matrix. It is also possible that general proteolytic activity by chymosin at this pH

Table 1
Analysis by CAT-ELISA of CAT released from immobilized fusion protein by chymosin treatment [a]

pH	Chymosin-treated immobilized StCL1CAT		Chymosin-treated immobilized StCL2CAT	
	Treatment time (h)			
	4	12	4	12
4.0	2	7.6	4.2	12
5.3	2	5.4	3.6	26.6
6.8	6.8	82.5	48.2 [b]	

The numbers represent the μg of CAT released.
[a] Fusion protein from 4 l of cell lysate was adsorbed on 0.5 ml of biotinylated beads, and after removal of nonspecifically adsorbed protein, the immobilized protein was treated with 10 U of chymosin.
[b] The specific CAT activity of this sample was 874 nmol min^{-1} mg^{-1} protein.

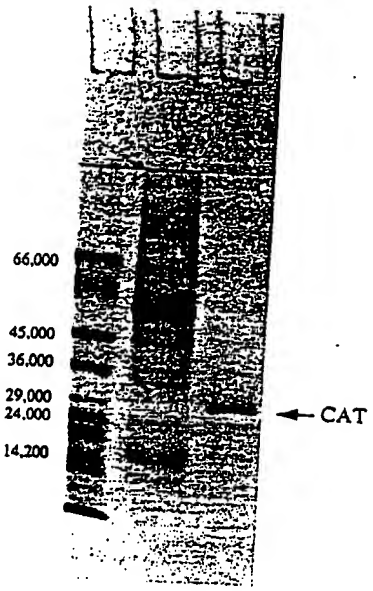


Fig. 4. SDS-PAGE analysis of chymosin-treated immobilized pStCL2CAT fusion protein. Lane 1, molecular mass markers; lane 2, crude protein from cells containing pStCL2CAT; lane 3, protein released from immobilized, chymosin-treated pStCL2CAT fusion protein.

inactivated CAT. In any case, CAT activity of protein released by chymosin could only be detected in the sample from pStCL2CAT transformed cells treated at pH 6.8. The specific activity of this protein was 874 U mg^{-1} protein. One unit of activity is defined as 1 nmol chloramphenicol acetylated per min. The optimum pH for CAT activity is pH 7.8; therefore, long incubations at the lower pH values may have inactivated CAT.

CAT released from immobilized protein from pStCL2CAT transformed cells treated with chymosin is indicated by the SDS-PAGE analysis shown in Fig. 4. The sample in lane three is chymosin-treated fusion protein obtained at pH 6.8. The major protein band, 30 kDa, corresponds to the calculated molecular mass of the released CAT portion of the fusion protein.

4. Conclusions

The chymosin-sensitive linker, CL2, encoding residues 97–113 of κ-casein, permitted rapid and

specific release of enzyme from streptavidin-CL2-enzyme fusion protein immobilized on biotinylated beads. In agreement with observations of chymosin activity on model peptide substrates, the 17-residue linker CL2 was cleaved more rapidly than the heptapeptide linker CL1. Combination of the chymosin-sensitive linker CL2 with the biotin-affinity streptavidin domain in fusion protein enzymes should permit simple production and purification of recombinant enzymes. The fusion protein can be purified by direct adsorption from crude cell lysates or culture media and the purified enzyme can be released by treatment with the readily available and inexpensive enzyme chymosin at pH near neutrality.

References

Dalgleish, D.G. (1982) The enzymatic coagulation of milk. In: Fox, P.F. (Ed.), Developments in Dairy Chemistry-1, Applied Science Publishers, London, pp. 157–187.

De Palma, A. (1994) Biotech advances in genetics and enzymology impact food processing. Genetic Engineering News 5, pp. 8, 37.

El-Negoumy, A.M. (1968) Starch gel electrophoresis of products of action of crystalline rennin on caseins and its components. J. Dairy Sci. 51, 1013–1017.

Fox, P.F. (1993) Exogenous enzymes in dairy technology-a review. J. Food Biochem. 17, 173–199.

Kumosinski, T.F., Brown, E.M., and Farrell, Jr., H.M. (1991) Three-dimensional molecular modeling of bovine caseins: κ-casein. J. Dairy Sci. 74, 2879–2887.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shaw, W.V. (1975) Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. Methods Enzymol. 43, 737–755.

Uhlen, M. and Moks, T. (1990) Gene fusions for purpose of expression: an introduction. Methods Enzymol. 185, 129–143.

Walsh, M.K. and Swaisgood, H.E. (1994) An *Escherichia coli* plasmid vector system for production of streptavidin fusion proteins: expression and bioselective adsorption of streptavidin-β-galactosidase. Biotechnol. Bioeng. 44, 1348–1345.

Foltmann, B. (1981) Gastric proteinases-structure, function, evolution and mechanism of action. Essays Biochem. 17, 52–84.

Keil, B. (1992) Specificity of Proteolysis. Springer-Verlay, Berlin. pp. 127–128.

Green, N.M. (1990) Avidin and streptavidin. Methods Enzymol. 184, 51–67.

Visser, S., van Rooijen, P.J., Schattenkerk, C. and Kerling, K.E.T. (1976) Kinetic studies of peptides of different chain length including parts of the sequence 101–112 of bovine κ-casein. Biochim. Biophys. Acta 443, 265–272.

Visser, S., van Rooijen, P.J. and Slangen, C.J. (1980) Peptide substrates for chymosin (rennin). Isolation and substrate behavior of two tryptic fragments of bovine κ-casein. Eur. J. Biochem. 108, 415–421.

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 7036–7040, July 1995
Genetics

Green fluorescent protein as a vital marker and reporter of gene expression in *Drosophila*

(targeted expression/enhancer trap/GAL4/confocal microscopy)

EDWARD YEH*, KERSTIN GUSTAFSON*, AND GABRIELLE L. BOULIANNE†

Centre for Research in Neurodegenerative Diseases, Department of Physiology and Zoology, Tanz Neuroscience Building, University of Toronto, Toronto, ON, Canada M5S 1A8

Communicated by Allan C. Spradling, Carnegie Institution of Washington, Baltimore, MD, April 21, 1995 (received for review March 16, 1995)

ABSTRACT **We have used the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* as a vital marker/reporter in *Drosophila melanogaster*. Transgenic flies were generated in which GFP was expressed under the transcriptional control of the yeast upstream activating sequence that is recognized by GAL4. These flies were crossed to several GAL4 enhancer trap lines, and expression of GFP was monitored in a variety of tissues during development using confocal microscopy. Here, we show that GFP could be detected in freshly dissected ovaries, imaginal discs, and the larval nervous system without prior fixation or the addition of substrates or antibodies. We also show that expression of GFP could be monitored in intact living embryos and larvae and in cultured egg chambers, allowing us to visualize dynamic changes in gene expression during real time.**

Development is the cumulative effect of dynamic changes in gene expression in different cells within an organism. At present, several techniques exist that allow an examination of gene expression through the measurement of either RNA or protein distribution within fixed tissue. Gene expression can be measured either directly by using probes and antibodies or indirectly by detecting the product of a fusion between the gene of interest and a reporter gene such as bacterial *lacZ* (1). In *Drosophila*, *lacZ* is often used in enhancer trap screens to identify genes that are expressed in a tissue-specific manner (2–4) or as a reporter to identify tissue-specific regulatory regions within known genes. All of these approaches are limited in that they only provide a static image of changes in gene expression during development. Furthermore, these techniques usually involve extensive manipulation including dissection of the tissue of interest, fixation, and the addition of various substrates or antibodies, and they are of limited use in living tissue.

Recently, Chalfie *et al.* (5) described the use of the green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* as a vital reporter for gene expression in both bacteria and *Caenorhabditis elegans*. In that study, GFP was placed under the transcriptional control of the *mec-7* promoter, which is activated in a small number of *C. elegans* neurons. GFP was nontoxic to cells, and its expression did not appear to interfere with cell growth and/or function. In addition, the green fluorescence did not appear to photobleach when viewed with fluorescein filter sets. These results suggested that GFP might be a powerful tool to examine changes in gene expression in living tissue.

Subsequently, GFP has been used in *Drosophila* to monitor the subcellular distribution of the exuperantia protein (exu) (6). In those studies, GFP was expressed as an in-frame fusion with the exu protein (encoded by *exu*) under the transcriptional control of its own promoter. The exu–GFP fusion protein was found to be expressed in the same pattern as the native exu protein. These results demonstrated the potential of GFP as a vital marker in *Drosophila*. However, the fact that GFP was produced as a exu–GFP fusion protein that was only expressed in adult ovaries precludes its use as a general marker/reporter gene. In theory, the exu–GFP fusion protein could be targeted to other tissues using various promoters, but the effects of ectopically expressing an exu–GFP fusion protein in other cell types are uncertain. Alternatively, additional fusion proteins could be generated with GFP, but this approach would be cumbersome, and the fusion products may be unstable, inactive, or nonfluorescent.

To make GFP more generally useful, we have utilized the GAL4 enhancer trap technique developed by Brand and Perrimon (7) to target expression of GFP. Here we show that GFP can be used as a vital marker of gene expression in a variety of living cell types at various developmental stages. We also demonstrate that GFP need not be expressed as a fusion protein in *Drosophila* but can be utilized directly as a reporter gene, much like *lacZ*. Finally, we show that GFP can be used to detect dynamic changes in gene expression in living tissue. Taken together, our results indicate that GFP will prove to be a powerful tool for viewing developmental changes within a living organism.

MATERIALS AND METHODS

Enhancer Detection Screen. GAL4-expressing enhancer trap lines were generated by mobilizing a single X chromosome-linked GAL4 *P*-element insertion (pGawB) as described (7). Four hundred crosses were set up to look for new insertion sites. In this study, only autosomal insertions were examined. Sixty-eight GAL4 insertion lines were obtained and balanced using standard genetic methods. Each of these was crossed to either Bg41-2 or Bg4-2-46 upstream activating sequence (UAS)–lacZ reporter lines, and β-galactosidase staining patterns were determined in embryos, imaginal discs, and ovaries. Out of 68 lines, 54 lines produced a detectable staining pattern. Lines that produced an interesting GAL4 expression pattern based on β-galactosidase staining were then crossed to a UAS–GFP line. The UAS–GFP line used in these experiments, GFPB1, contains a homozygous viable insertion on chromosome 3.

UAS–GFP Construct. The GFP coding region was isolated from plasmid TU#65, which contains the GFP cDNA in pBS(+) (Stratagene) as a *Kpn* I–*Eco*RI fragment and subcloned into pUAST. pUAST is a *P*-element vector based on pCaSpeR3 containing five optimal GAL4 binding sequences followed by a multiple cloning site (7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GFP, green fluorescent protein; UAS, upstream activating sequence; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside.

*E.Y. and K.G. contributed equally to the research presented in this paper.
†To whom reprint requests should be addressed.

Transformations. Transgenic flies carrying the UAS–GFP construct were generated by injecting pUAS–GFP DNA at a concentration of 400 μg/ml with the helper plasmid pπ25.7wc at a concentration of 100 μg/ml into embryos of the w^1 strain (8, 9) using standard methods (10). A total of eight different lines were generated with the UAS–GFP insertion on X chromosome, chromosome 2, or chromosome 3.

Detection of *lacZ* by Antibody Immunocytochemistry and 5-Bromo-4-chloro-3-indolyl β-D-galactoside (X-Gal) Staining. The procedures used to collect embryos, remove the vitelline membrane, and stain whole-mount embryos with antibodies are described (11). The primary antibody, an IgG fraction rabbit anti-β-galactosidase from Cappel, was used at a concentration of 1:4000. The secondary antibody, a horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad), was used at 1:200. Imaginal discs were dissected in chilled phosphate-buffered saline (PBS) and fixed for 20 min in 0.75% glutaraldehyde. After being washed once with PBS/0.1% Triton X-100, cells were stained as described (12) and mounted in 70% glycerol. Ovaries were dissected and fixed in 1% glutaraldehyde for 20 min. Fixative was removed, and the ovaries were stained as described (13). After being washed in PBS/0.1% Triton X-100, they were mounted in 98% glycerol. Photographs of β-galactosidase histochemistry and immunocytochemistry were taken on a Nikon Optiphot 2 microscope with Nomarski optics, using a Nikon FX-35 camera.

Visualization of GFP Using Confocal Microscopy. Confocal images of GFP expression were taken on either a Leica DM IRB inverted laser confocal microscope using a standard fluorescein isothiocyanate filter providing excitation at 490 nm and emission at 527 nm or a Bio-Rad model MRC600 using a standard fluorescein isothiocyanate filter on a Nikon Optiphot 2 microscope. In all cases, image files were processed using a computer-based graphic system (Corel 4.0) where they were arranged and annotated. Images acquired of imaginal discs, the larval nervous system, and developing egg chambers were not further processed. All other images were processed to adjust the brightness and contrast of the image using Corel PHOTOPAINT. Imaginal discs from F_1 larvae resulting from a GAL4 line/UAS–GFP cross were dissected in distilled water and mounted immediately in 70% (vol/vol) glycerol/30% 0.1 M Tris (pH = 9). Larval central nervous systems were dissected in Schneiders medium according to standard procedures (14). F_1 embryos were dechorionated with 3% sodium hypochlorite, rinsed with distilled water, and mounted in 70% glycerol/30% 0.1 M Tris (pH = 9.0). Ovaries were dissected and mounted in PBS from 2-day-old F_1 virgin females. No fixatives were used in any of these preparations. Images of GFP expression in developing ovaries were derived as follows. Ovaries were dissected in Schneider's medium/10% fetal calf serum. Stage-8 egg chambers were dissected out of the epithelial sheath overlaying ovarioles and transferred to a microscope slide with medium. An artificial well was created on the slide using stacked slips of paper covered with vacuum grease to hold enough medium to bathe the sample and to support a coverslip. A *z*-series of confocal images was then obtained every hour for a total of 4 hr.

RESULTS

To determine whether GFP could be used as a vital marker/reporter in a variety of tissues during *Drosophila* development we generated transgenic lines containing the GFP cDNA under the transcriptional control of the yeast UAS. Three independent lines were tested by crossing them to several GAL4 enhancer trap lines that we had generated according to the protocol outlined in Brand and Perrimon (7). No apparent difference was observed in the ability to detect GFP from any of the three lines tested. The results presented here were obtained with the GFP-B1 line, which is a homozygous viable insertion of the UAS–GFP transgene on chromosome 3.

Expression of GFP was first examined by crossing the UAS–GFP-B1 line to three GAL4 enhancer trap lines that are expressed within adult ovaries. In all cases, appropriate expression of the UAS–GFP transgene was confirmed by comparing the results obtained with histochemical results of parallel crosses of the GAL4 lines with a UAS–lacZ line (Fig. 1 *A*, *C*, and *E*). The GAL4 lines used in these experiments targeted expression of GFP to posterior follicle cells (Fig. 1*B*), stalk cells (Fig. 1*D*), and nurse-cell-associated follicle cells (Fig. 1*F*). Expression of GFP could be detected in both freshly dissected ovaries and in fixed tissue (data not shown). By using the GFP marker in combination with confocal microscopy, we obtained greater resolution of the expression pattern than observed using X-Gal staining. The apparent reduction in the number of cells that express GFP in the posterior and nurse-cell-associated follicle cells is due to the optical sectioning of the confocal microscope (Fig. 1 *B* and *F*). Thus, GFP can be expressed and detected in ovaries not only as a fusion protein (6) but as a reporter gene as well. As previously noted (6), we could also detect minimal levels of autofluorescence within late egg chambers. However, using a barrier filter with a wavelength cut-off of 580 nm, we could distinguish between GFP and autofluorescence: GFP emits light maximally at 509 nm and is not detectable under these conditions, whereas autofluorescence can still be observed.

We also examined whether GFP could be detected in larval tissues by crossing the UAS–GFP-B1 line to several GAL4

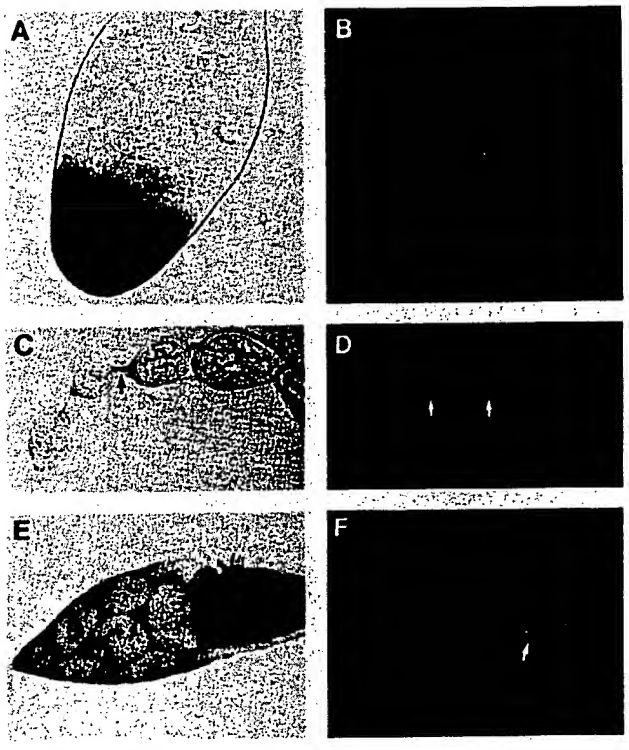


FIG. 1. Detection of GFP during oogenesis. Three GAL4 enhancer trap lines were used to detect expression of GFP during oogenesis. Expression of GFP was confirmed by crossing each GAL4 line to a UAS–lacZ line and staining ovaries with X-Gal. *A*, *C*, and *E* represent X-Gal staining; *B*, *D*, and *F* represent the same GAL4 lines examined for GFP expression. (*A* and *B*) GAL4 line A62, which directs expression of β-galactosidase and GFP in posterior follicle cells. (*C* and *D*) GAL4 line A39, which directs expression of β-galactosidase and GFP to stalk cells (see arrows). (*E* and *F*) GAL4 line A90, which directs expression of β-galactosidase and GFP to nurse-cell-associated follicle cells (see arrows). (*A–D*, ×94; *E*, ×100.)

enhancer trap lines that target GFP expression to specific cells in imaginal discs and the larval nervous system (Fig. 2). Illustrated are examples where GFP expression is targeted to the optic lobe (Fig. 2*B*), the eye disc and optic lobe (Fig. 2*D*), and a wing imaginal disc (Fig. 2*F*). Interestingly, when GFP was expressed in optic lobe neurons, the protein was not restricted to cell bodies within the larval nervous system but could also be detected in their processes. As seen in Fig. 2*D*, GFP clearly marked photoreceptor axons as they projected from the eye disc to innervate the optic lobe.

To further examine whether GFP could be used to identify nerve terminals, we used GAL4 enhancer trap lines to target expression of GFP to motoneurons within the larval nervous system (Fig. 3). These experiments clearly demonstrated that GFP could be detected both in neuronal cell bodies and in the processes immediately extending from the cell bodies (Fig. 3*A*). Similar to that observed in *C. elegans* (5), GFP could also be detected within nerve terminals at the point where they innervated specific muscles. For example, in a nerve terminal that innervates muscle 12 of the larval abdomen (Fig. 3*B*), GFP clearly outlines both the preterminal region and synaptic boutons (Fig. 3*B*). The only significant photobleaching observed was within the nerve terminals and seen only after prolonged exposure to the laser beam. However, fluorescence was recovered after a brief rest period in the absence of the laser beam. In the larval tissues examined, no autofluorescence was observed, and it was necessary to artificially increase the background by adjusting the baseline fluorescence using the black level control. This adjustment increased the total brightness of the image by a constant and allowed us to visualize the underlying structures for photography.

These experiments clearly demonstrate that GFP can be used to detect gene expression in a variety of freshly dissected tissues in *Drosophila* without any requirement for fixation or additional substrates. To determine whether GFP can also be used in intact living animals we examined GFP expression in embryos and larva (Fig. 4). Embryos were dechorionated, mounted in halocarbon oil on a glass slide, and viewed by confocal microscopy. Exposure of the embryo to the laser beam for the short periods of time required to obtain an image

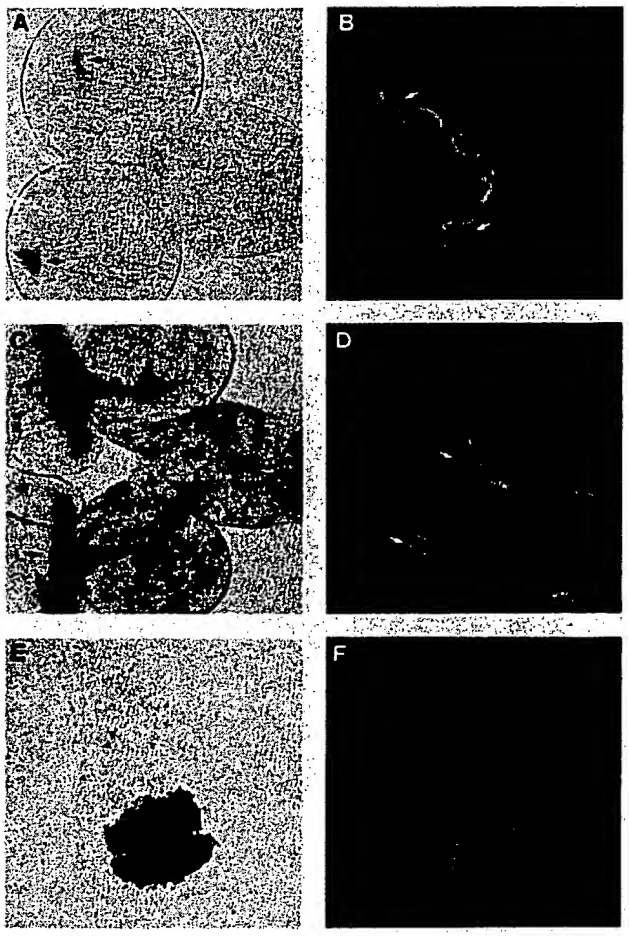


FIG. 2. Detection of GFP in larval imaginal discs and central nervous system. Three GAL4 enhancer trap lines were used to detect expression of GFP in larval imaginal discs and the central nervous system. To confirm that GFP was expressed in the appropriate pattern, each GAL4 line was crossed to a UAS–lacZ line and examined by X-Gal staining. *A*, *C*, and *E* represent the X-Gal staining pattern for each GAL4 line; *B*, *D*, and *F* represent the GFP expression pattern. (*A* and *B*) GAL4 line A95, which directs expression of β-galactosidase and GFP to neuronal cell bodies within the optic lobe and to their processes that extend from one lobe to the other (see arrows). The optic lobes are located at left and the ventral ganglia are located at right. In this preparation, the eye discs have been removed. (*C* and *D*) GAL4 line B41, which directs expression of β-galactosidase and GFP to photoreceptor neurons within the eye disc and to their processes, which innervate deep within the optic lobe. Arrows indicate location of the photoreceptor cells within the eye disc. (*E* and *F*) GAL4 line C5, which directs expression of β-galactosidase and GFP to the region of the wing imaginal disc, which will give rise to the wing blade. (*A–F*, ×110.)

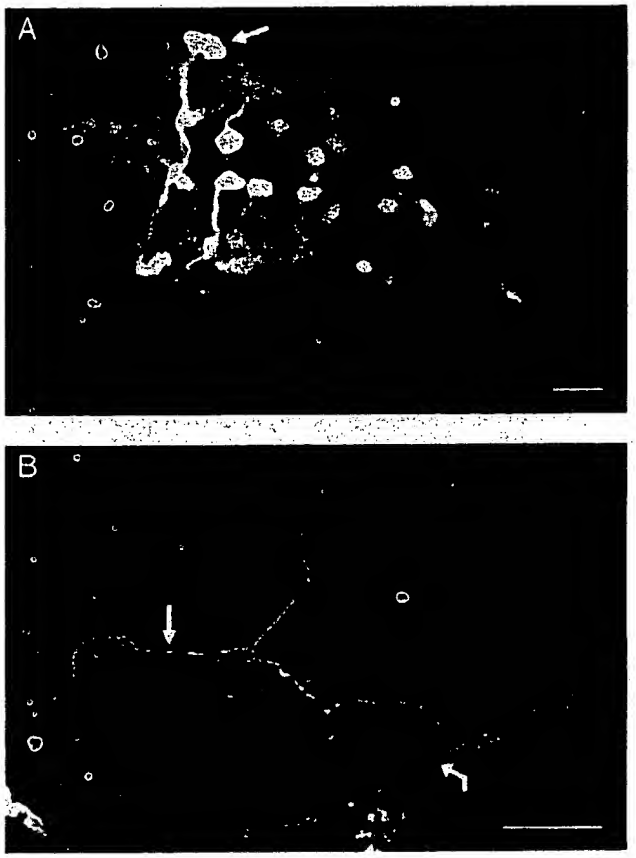


FIG. 3. Detection of GFP in motoneuron cell bodies and nerve terminals. The GAL4 enhancer trap line D42 was used to target expression of GFP to motoneurons within a living third-instar larvae. (*A*) Expression of GFP can be detected within motoneuron cell bodies and in the processes that immediately extend from them. The arrow points to a specific motoneuron within the larval CNS. (*B*) Expression of GFP can also be detected at the nerve terminal and within synaptic boutons (straight arrow). Autofluorescence can also be detected within the trachae (bent arrow). (Bar = 50 μm.)

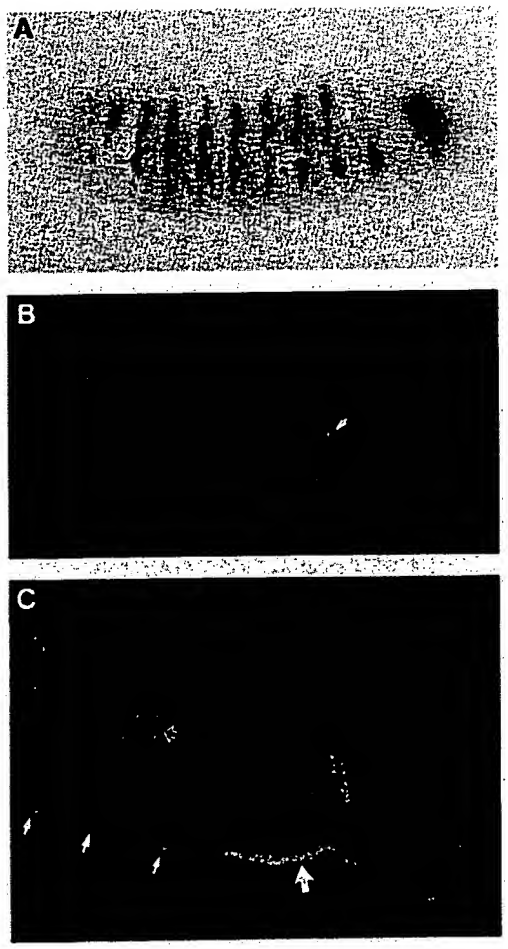


FIG. 4. Detection of GFP in living embryos and larvae. Expression of GFP can also be detected in embryos or larvae without any prior dissection or fixation. (*A*) Anti-β-galactosidase staining of embryos expressing β-galactosidase driven from the GAL4 line C41. Expression can be detected in most, but not all, of the cells which comprise the peripheral nervous system of the embryo. The arrow points to a group of cells within the lateral cluster of the peripheral nervous system. (*B*) Expression of GFP driven by the GAL4 line C41 can be detected within the peripheral nervous system of live embryos. (*C*) Expression of GFP driven by the GAL4 line C38 can be detected within the larval salivary glands (large filled arrow) and pair of cells along the body wall within each segment (small filled arrows). Autofluorescence from the gut is indicated by an open arrow. (*A–C*, ×55.)

(<1 min) did not alter its viability, and embryos were observed to hatch into larvae (data not shown). Expression of GFP is shown for one GAL4 line, C41, which is expressed in much of the embryonic peripheral nervous system (Fig. 4*B*). Appropriate expression of GFP was confirmed by immunostaining embryos with an anti-β-galactosidase antibody (Fig. 4*A*). While all of the peripheral neurons cannot be observed in a single confocal optical section, the overall pattern is maintained. Similar observations can also be made in intact larvae. Fig. 4*C* shows a GAL4 enhancer trap line that targets GFP to both salivary glands and pairs of cells that are distributed segmentally along the larval body wall. In larvae, as in embryos, development was not affected by exposure to the laser beam. In contrast to larval imaginal discs and nervous system, whole embryos and larvae have significant levels of autofluorescence due to the yolk and gut, respectively. These levels, however, could be resolved by using appropriate barrier filters as described above.

To determine whether GFP could be used to detect dynamic changes in gene expression within living tissues in real time, we

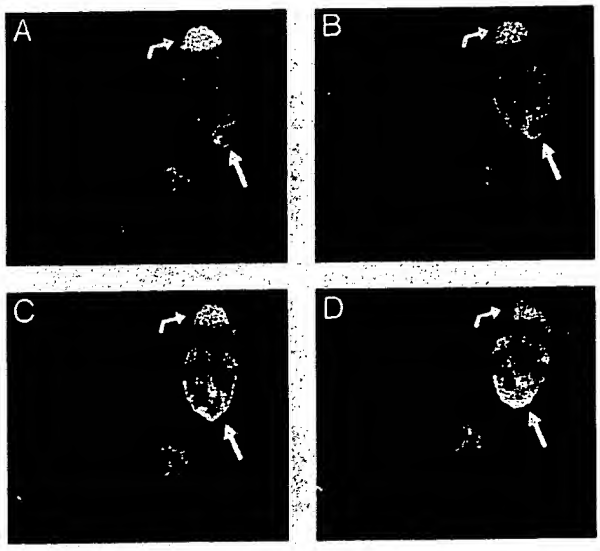


FIG. 5. Changes in GFP expression can be detected in cultured egg chambers. Expression of GFP was monitored in cultured stage-8 egg chambers during a 4-hr period. A *z*-series of confocal images was obtained every hour for the entire 4-hr period. Images represent a particular focal plane from each *z*-series. Expression of GFP in egg chambers is directed by GAL4 line A90, which targets expression to nurse-cell-associated follicle cells. (*A*) Detection of GFP after 1 hr in culture. Little to no expression of GFP can be detected within follicle cells (straight arrow). Low levels of autofluorescence, however, can be observed within the oocyte (bent arrow). (*B*) GFP expression begins to be detected after 2 hr in culture in follicle cells at the anterior end of the stage-8 egg chamber (straight arrow). In contrast, the level of autofluorescence from the oocyte decreases (bent arrow), and no changes are observed in earlier egg chambers. (*C*) After 3 hr, levels of GFP expression within nurse-cell-associated follicle cells increase (straight arrow). (*D*) After 4 hr, GFP expression can be detected in all nurse-cell-associated follicle cells and is particularly high in follicle cells at the anterior tip of the egg chamber (straight arrow). No changes in fluorescence are detected in earlier egg chambers, and the autofluorescence in the oocyte remained low (bent arrow). (*A–D*, ×90.)

performed time-lapse confocal microscopy of developing egg chambers (Fig. 5). These experiments were done by using a GAL4 enhancer trap line, A90, which targets expression of UAS–GFP to nurse-cell-associated follicle cells (Fig. 1*C*). Egg chambers that were dissected and cultured in Schneiders medium over a period of 4 hr are shown. Initially, no GFP could be detected within stage-8 egg chambers (Fig. 5*A*). However, by 1 hr, GFP expression was observed at the anterior end of the stage-8 egg chamber, and this expression increased steadily over time (Fig. 5 *B–D*). Expression of GFP is specific to nurse-cell-associated follicle cells and restricted to stage-8 egg chambers. No GFP could be detected in stage-2 to -7 egg chambers. These results clearly show that GFP can be used as a reporter to monitor activation of gene expression in living tissue over time. Whether GFP can also be used to monitor cessation of gene expression remains to be determined and will depend on the stability of GFP in various cell types during development.

DISCUSSION

The ability to study development as it occurs within an organism relies on the availability of techniques that can detect changes in gene expression within specific cells or tissues during cell movements and migrations. We have used the GFP from the jellyfish, *A. victoria*, as a viable marker in *Drosophila* to observe such changes within living tissues. GFP was expressed as a nonfused protein under the transcriptional control of a yeast UAS and targeted to specific cell types during

development by crossing to a variety of GAL4-expressing enhancer trap lines. These studies clearly show that GFP can be used as a reporter gene, much like bacterial *lacZ*, to detect expression of specific genes in a variety of cell types during *Drosophila* development. However, in contrast to *lacZ*, GFP can be visualized in live tissue without fixation or addition of specific substrates and often without any dissection. This result permits the monitoring of gene expression within a living organism over time.

The ability to use GFP as a vital marker/reporter in *Drosophila* suggests a number of other interesting applications. For example, development of an enhancer trap vector system based on GFP suggests the possibility of bulk screening, whereby specific expression patterns could be detected in embryos in the F_1 generation. Because GFP can function as a reporter gene, it should be able to replace *lacZ* in other assays, such as promoter mapping. In addition, it may be possible to use GFP to sort pure populations of live cells using a fluorescence-activated cell sorter as originally described by Krasnow *et al.* (15) using a fluorogenic β-galactosidase substrate. GFP could also be recombined onto various balancer chromosomes to allow for rapid identification of embryos/larvae containing the specific balancer, much like *lacZ*-marked balancers that are currently available. GFP balancers would permit identification and selection of homozygous mutants based on the absence of GFP expression. *P*-element vectors could also be generated using GFP as a reporter gene and used to characterize various mutant phenotypes arising from *P*-element insertional mutagenesis. Because GFP is expressed not only in cell bodies but also in processes, this may be particularly useful in identifying axon guidance or pathfinding mutants. Until recently, these mutants were difficult to identify, as β-galactosidase fails to readily diffuse into axons. Alternative approaches in which β-galactosidase is expressed as a kinesin–lacZ (16) or Tau–lacZ fusion protein (17) are more efficient at detecting axonal processes but still require fixation and the addition of specific substrates or antibodies to detect expression, which only provides a static image of the axonal process. The expression of GFP in nerve terminals may also be useful for studying synaptogenesis. For example, GAL4 lines that target expression of GFP to photoreceptor axons could be used to examine their ability to form appropriate synapses within the optic lobe. Expression of GFP in nerve terminals could also be used to identify specific neurons and/or synaptic boutons, which could then be analyzed electrophysiologically.

Finally, we have shown that GFP can be used to detect changes in gene expression in living tissue. This result suggests that GFP could be used in fate mapping or lineage analysis experiments. GFP could also be used to monitor changes in cell migrations or cell shape such as occur during germ-band extension (18) and the migration of pole cells (19), follicle cells (13), and tracheal cells (20, 21) during development. The identification of mutations that affect the pattern of specific cell migrations combined with the ability to visualize the cells in living tissues using GFP should provide insight into the mechanisms that control cell movements as they occur within the organism.

We thank Drs. A. Brand and N. Perrimon for kindly providing us with the pUAST vector, the GAL4–lethal/FM7 jumpstarter and UAS–lacZ strains and Dr. M. Chalfie for the GFP cDNA. We thank L. Zhou for providing us with excellent technical help during the course of these experiments and B. Stewart for help with larval central nervous system preparations. We are indebted to J. Wadia, J. Georgiou, and Drs. W. G. Tatton and M. Charlton for excellent advice and help with all aspects of confocal microscopy. The Bio-Rad MRC600 confocal microscope facility was provided by the Neuroscience Network. We are most grateful to Drs. W. S. Trimble, H. Krause, and J. Culotti for critical comments on the manuscript. This work was supported by the National Science and Engineering Council of Canada, the Medical Research Council of Canada, and the Amyotrophic Lateral Sclerosis Society of Canada. E.Y. is partially supported by a University of Toronto Open Scholarship. G.L.B. is supported by the Alzheimers Association of Ontario.

1. Silhavy, T. J. & Beckwith, J. R. (1985) *Microbiol. Rev.* **49**, 398–418.
2. O'Kane, C. J. & Gehring, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9123–9127.
3. Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. & Gehring, W. J. (1989) *Genes Dev.* **3**, 1288–1300.
4. Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Caretto, R., Uemura, T., Grell, E., Jan, L. Y. & Jan, Y. N. (1989) *Genes Dev.* **2**, 1273–1287.
5. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. (1994) *Science* **263**, 802–805.
6. Wang, S. & Hazelrigg, T. (1994) *Nature (London)* **369**, 400–403.
7. Brand, A. H. & Perrimon, N. (1993) *Development (Cambridge, U.K.)* **118**, 401–415.
8. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218**, 348–353.
9. Karess, R. E. & Rubin, G. M. (1984) *Cell* **38**, 135–146.
10. Spradling, A. C. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford), pp. 175–197.
11. Bodmer, R. & Jan, Y. N. (1987) *Roux Arch. Dev. Biol.* **196**, 69–77.
12. Wagner-Bernholz, J. T., Wilson, C., Gibson, G., Schuh, R. & Gehring, W. J. (1991) *Genes Dev.* **5**, 2467–2480.
13. Montell, D. J., Rorth, P. & Spradling, A. C. (1992) *Cell* **71**, 51–62.
14. Budnick, V., Zhong, Y. & Wu, C.-F. (1990) *J. Neurosci.* **10**, 3754–3768.
15. Krasnow, M. A., Cumberledge, S., Manning, G., Herzenberg, L. A. & Nolan, G. P. (1991) *Science* **251**, 81–85.
16. Giniger, E., Wells, W., Jan, L. Y. & Jan, Y. N. (1993) *Roux Arch. Dev. Biol.* **202**, 112–122.
17. Callahan, C. A. & Thomas, J. B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5972–5976.
18. Irvine, K. D. & Wieschaus, E. (1994) *Development (Cambridge, U.K.)* **120**, 827–841.
19. Jaglarz, M. K. & Howard, K. R. (1994) *Development (Cambridge, U.K.)* **120**, 83–89.
20. Klambt, C., Glazer, L. & Shilo, B.-Z. (1992) *Genes Dev.* **6**, 1668–1678.
21. Reichman-Fried, M., Dickson, B., Hafen, E. & Shilo, B.-Z. (1994) *Genes Dev.* **8**, 428–439.

Theor Appl Genet (2002) 105:878–889
DOI 10.1007/s00122-002-1039-5

P. Vain · V. A. James · B. Worland · J. W. Snape

Transgene behaviour across two generations in a large random population of transgenic rice plants produced by particle bombardment

Received: 17 November 2001 / Accepted: 20 January 2002 / Published online: 22 August 2002

Abstract The relationship between transgene copy number, rearrangement levels, inheritance patterns, expression levels, transgene stability and plant fertility was analysed in a random population of 95 independently transformed rice plant lines. This analysis has been conducted for both the selectable marker gene (*aphIV*) and the unselected reporter gene (*gusA*), in the presence or absence of flanking Matrix Attachment Regions (MARs) in order to develop a better understanding of transgene behaviour in a population of transgenic rice plants created by particle bombardment. In the first generation (T_0), all the independently transformed plant lines contained and expressed the *aphIV* gene conferring resistance to hygromycin, but only 87% of the lines were co-transformed with the unselected *gusA* marker gene. Both transgenes seemed to be expressed independently. Most lines exhibited complex transgene rearrangements as well as an intact transgene expression unit for both *aphIV* and *gusA* transgenes. Transgene copy number was proportional to the quantity of DNA used during bombardment. In T_0 plants, high *gusA* copy number significantly decreased GUS expression levels but there was no correlation between expression level and transgene copy number across the entire population of lines. Four main factors impaired transgene expression in primary transgenic plants (T_0) and their progeny (T_1): (1) absence of transgene expression in T_0 plants (41% of lines), (2) sterility of T_0 plants (28% of lines), (3) non-transmission of intact transgenes to some or all progenies (at least 14% of lines), and (4) silencing of transgene expression in progeny plants (10% of lines). Transgene stability was significantly related to differences in transgene structure and expression levels. The presence of Rb7 MARs flanking the *gusA* expression unit had no effect on plant fertility or non-transmission of transgenes, but provided copy number-dependent expression of the transgene and improved expression levels and stability over two generations. Overall, only 7% of the plant lines without MARs and 17% of the lines with MARs initially generated, exhibited stable transgene expression over two generations.

Keywords Transgene expression · *Oryza sativa* · Particle bombardment · Fertility · Matrix · Attachment Regions

Communicated by C. Möllers

P. Vain (✉) · V.A. James · B. Worland · J.W. Snape
John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
e-mail: philippe.vain@bbsrc.ac.uk
Tel.: +44-1603-450612, Fax: +44-1603-450023

Introduction

In the past ten years there has been great progress in cereal transformation technologies. However, transgene expression in plants remains largely unpredictable, and there is considerable variation in expression levels and stability between independently transformed plants (Jones et al. 1985; Peach and Velten 1991; Walters et al. 1992). Different integration sites, copy numbers and transgenic locus configurations, as well as epigenetic silencing mechanisms, can all contribute to this variability (reviewed by Finnegan and McElroy 1994; Meyer 1995; Matzke and Matzke 1998; Iyer et al. 2000). Experimental procedures such as transformation systems (*Agrobacterium* vs direct transfer of DNA), construct configuration (Breyne et al. 1992), promoters (Mlynárová et al. 1995), coding sequences, terminators, selection strategy (Bhattacharyya et al. 1994), flanking Matrix Attachment Regions (MARs) (Mlynárová et al. 1994) or the plant tissue analysed (Ülker et al. 1999) have also been reported to influence transgene structure or expression in plants. The multiplicity of these factors, and their interactions, contribute strongly to the unpredictability, variability and instability of transgene expression in plants. This problem is particularly acute in plants generated by direct transfer of DNA (electroporation of cells and protoplasts, particle gun bombardment, silicon carbide fibres) due to the complex transgenic loci created in the plant genome. The numerous and uncontrolled transgene rearrangements, high gene copy number and systematic

transgene linkage (Lyznik et al. 1989; Gordon-Kamm et al. 1990; Wan and Lemaux 1994; Pawlowski and Somers 1996) can all favour variable and unstable transgene expression (Finnegan and McElroy 1994; Hansen and Chilton 1996; Matzke and Matzke 1998). Until recently most transgenic cereal crops were produced by technologies based on direct transfer of DNA, particularly particle gun bombardment (Christou 1996). In transgenic cereals, more than 50% of specific transgenes can be inactivated over successive generations (Pawlowski and Somers 1996; Iyer et al. 2000). High levels of transgene expression can also be hard to achieve in specific cereal tissues. These problems make molecular genetic studies difficult, and frustrate attempts at crop improvement through genetic engineering. Additionally, they create difficulties in predicting transgene behaviour when transgenes are transferred by conventional crossing, and in predicting gene flow by accidental out-crossing.

Numerous transgenic studies have been conducted in cereals; however, the relationship between transgene structure, expression level, inheritance pattern and stability in populations of transgenic plants over several generations often remains unclear. This is due to the difficulties associated with developing combined analysis of all these factors in large populations of transgenic plants over several generations. In theory, such a multi-factorial study would require most of the following: (1) production of large and random populations of independent transformation events (cell/callus lines), (2) regeneration, without phenotypic selection, of several primary plants (T_0) per transformation event, (3) study of all plants regenerated from each transformation event, expressing or not expressing the transgene(s), (4) detailed study of transgene structure (copy number, integration pattern, integration site), (5) quantification of transgene expression level during plant development using standardised conditions and protocols (controlled environment, defined plant developmental stages and tissue sampled), (6) assessment of plant development characteristics (growth, fertility), (7) study of transgene inheritance at the structural level, (8) study of transgene inheritance at the expression level, (9) study of transgene structure in the progeny, (10) quantification of transgene expression levels in the progeny, and (11) statistical analysis of these factors and their interactions. To-date, transgenic studies in cereals have addressed most of these aspects individually but rarely in combination. The absence of such multi-factorial analysis in a large and random population of transgenic plants, over generations, limits our understanding of which factors are the most significant for transgene behaviour and how these factors interact. It also limits the possibility of identifying sub-populations of plant lines with specific characteristics and behaviour. Most importantly it limits our ability to predict transgene behaviour across generations.

In the present study we have analysed the relationship, over two generations, between transgene copy number, rearrangement levels, inheritance patterns, expression level and stability for both the *aphIV* selectable marker gene and for the unselected *gusA* reporter gene, as well as plant fertility in a large random population of transgenic rice plants created by particle bombardment. This analysis has provided a better understanding of transgene behaviour in a population of transgenic plants as well as some understanding of how transgene structure and expression level can influence aspects of plant development and transgene stability.

Materials and methods

Rice transformation procedures

African elite rice (*Oryza sativa* L.) variety ITA212 was co-transformed by particle gun bombardment with the plasmid pJIC201 (ubi-5′ region :: *aphIV* :: SoyT) and one of the following plasmids: pGHNC12 (CaMV35S :: *gusA* :: nosT), or pGHNC11 (Rb7MAR :: CaMV35S :: *gusA* :: nosT :: Rb7MAR), or pGA984 (ARS1MAR :: CaMV35S :: *gusA* :: nosT :: ARS1MAR), as previously described (Vain et al. 1999). Independently transformed rice callus lines were selected for hygromycin resistance. Five transgenic T_0 plants were regenerated from each callus line (i.e. a transformation event) representing an independent plant line. Transgenic plants were transferred to a controlled environment room for growth to maturity. All transgenic plants produced were used in further experiments to ensure the study of randomised independent transformation events with the widest spectrum of expression for the non-selected *gusA* gene.

Analysis of GUS activity

Fluorometric analysis for β-glucuronidase activity was carried out on leaf tissue from rice plants at the five-leaf stage, following the method of Jefferson (1987). Fluorescence was measured using a Titertek Fluoroskan II after 0-, 30- and 60-min incubation. Each assay was performed in triplicate. Protein content was determined using a Bio-Rad protein assay kit. Data were expressed as pmol of 4-methylumbelliferone (MU) min^{-1} mg^{-1} of extracted protein. GUS activity was measured in five different T_0 plants and eight different T_1 plants for each independent line. The background activity (33 ± 4 pmol MU min^{-1} mg^{-1} protein) was subtracted from all fluorometric GUS measurements as previously described (Vain et al. 1999).

Detection of transgenic plants by the Polymerase Chain Reaction (PCR)

DNA was isolated from rice plants and PCR reactions were carried out as previously described (Vain et al. 1998). Three primer sets were used on each DNA sample: (1) one to amplify the 1,200-bp single-copy rice RFLP probe C213 (forward: 5′-AAAGGACCGGAATGACCACAA-3′; reverse: 5′-GAATGAACCACGCCCAAGAGT-3′) in order to ensure that each DNA sample was suitable for PCR amplification, (2) another to amplify a 1,271-bp fragment containing the *aphIV* gene (forward: 5′-ACTCACCGCGACG TCTGTCG-3′; reverse: 5′-GATCTCCAATCTGCGGGATC-3′), (3) the other to amplify a 2,038-bp fragment of the CaMV35S :: *gusA* expression cassette (forward: 5′-CCCACCCACGAGGAGCAT-3′; reverse: 5′-GCGCCAGGAGAGTTGTTGATT-3′). Additional PCR primers nested within the *aphIV* or *gusA* expression cassettes were used to amplify smaller fragments within these regions.

Gene copy number analysis

Genomic DNA extraction and Southern analyses were performed on primary transformed rice plants (T_0) as previously described

(Vain et al. 1999). Membranes were hybridised with probes generated by PCR amplification of 701 nt of the *gusA* gene or 981 nt of the *aphIV* gene, or with the R2272 rice RFLP probe (to control DNA loading). The filters were analysed by autoradiography followed by densitometry (Vain et al. 1999). Final copy number was calculated by linear regression analysis based upon hybridisation signals obtained from the reconstitution standards and normalisation by DNA loading (Vain et al. 1999).

Statistical analysis

Statistical analyses, following the requirements of each test, were performed using Minitab 13.1 and Genstat 5 software. The normality of distribution of each data set was evaluated employing the Anderson Darling test. Variances were compared using the Levene's test. Data sets were compared using ANOVA. Data sets not meeting ANOVA requirements were compared using the non-parametric Kruskal-Wallis and Mann-Whitney tests. Linear regression analysis was only performed on normally distributed data sets.

Transgene inheritance study

T_1 seeds were obtained by self-pollination of primary transformed rice (T_0) plants.

Segregation analysis at the structural level was conducted by germinating seeds on MSR6 medium (Vain et al. 1998) without hygromycin and PCR analyses were performed to test for the presence of the *aphIV* and *gusA* transgenes in up to 40 random T_1 seedlings from each T_0 plant. When PCR reactions were negative for the *aphIV* and *gusA* transgenes, PCR analysis was conducted for the presence of the C213 RFLP probe to confirm that DNA extractions were suitable for PCR amplification.

Segregation analysis at the expression level was assessed qualitatively by histochemical GUS staining (Jefferson et al. 1987) of the T_1 seed endosperm and germination of the corresponding isolated embryo on hygromycin-containing medium (MSR6H50, Vain et al. 1998). All plant lines exhibiting skewed segregation at the expression level were re-analysed at the structural level as described above.

Results and discussion

Transgene structure in T_0 rice plants

Transformed rice plants were regenerated from 95 independent transgenic callus lines co-bombarded with a plasmid containing the *aphIV* hygromycin resistance gene (PJIC201) and a plasmid containing the *gusA* reporter gene, either as a simple expression cassette (pGHNC12) or flanked by the Rb7 MARs from tobacco (pGHNC11) or by the ARS1 MARs from yeast (pGA984). PCR and Southern analyses showed that all the transgenic plant lines contained the *aphIV* gene, but that only 87% of the lines (83 out of 95 lines) contained the unselected *gusA* reporter gene (Table 1). Such a high co-transformation frequency is common in transgenic plants transformed by direct transfer of DNA (Lyznik et al. 1989; Gordon-Kamm et al. 1990). It may result from

Table 1 Transgene copy number and expression in T_0 and T_1 rice plants. *n*: number of independently transformed plant lines. *ne*: number of plant lines expressing the *gusA* gene. *nf*: number of fertile plant lines (expressing and not expressing the *gusA* gene). *nfh*: number of fertile plant lines which transmit intact transgenes to subsequent T_1 generation (expressing and not expressing the *gusA* gene). + in pmol of MU min^{-1} mg^{-1} of extracted protein. Mean: row entries followed by different letters (a/b) are significantly different at $P < 0.05$ by ANOVA. Data sets not meeting ANOVA requirements were analysed using Kruskal-Wallis and Mann-Whitney tests. CV (Coefficient of variation): row entries followed by the same letter are not significantly different at $P > 0.05$ using Levene's test performed on raw data expressed as a percentage of the mean

1st Plasmid (P1): 2nd Plasmid (P2):		PJIC201 (*aphIV*) pGHNC12 (*gusA*)	PJIC201 (*aphIV*) pGHNC11 (Rb7-*gusA*-Rb7)	PJIC201 (*aphIV*) pGA984 (ARS1-*gusA*-ARS1)
Lines with P1 only	*n*	3	3	6
Lines with P1 + P2	*n*	29	29	25
gusA copy number	*n*	29	29	25
	Mean	34 a	24 ab	15 b
	CV	84 a	95 a	74 a
aphIV copy number	*n*	29	29	25
	Mean	6 a	10 a	8 a
	CV	69 a	73 a	67 a
GUS expression levels (T_0) +	*n*	29	29	25
	Mean	1,158 a	2,935 b	3,548 b
	CV	141 a	93 a	104 a
	ne	17	24	18
	Mean	1,964 a	3,545 ab	4,924 b
	CV	88 a	73 a	71 a
Fertility (T_0)	*nf*	16	17	18
(# of seeds)	Mean	84 a	97 a	116 a
	nfh	12	11	11
	Mean	107 a	73 a	146 a
	CV	98 a	58 a	77 a
GUS expression levels (T_1) +	*n*	5	9	7
	Mean	1,711 a	2,801 a	3,829 a
	CV	96 a	88 a	82 a
Fertility (T_1)	*nfh*	12	11	11
(# of seeds)	Mean	144 a	167 a	174 a
	CV	97 a	65 a	77 a

either extra-chromosomal recombination before integration, duplication of the integrated transgenes or multiple adjacent integration sites. The random population of 83 co-transformed lines was composed of the following: (1) 29 independent plant lines without MARs (PGHNC12), (2) 29 independent plant lines containing Rb7 MARs (pGHNC11), and (3) 25 independent plant lines containing ARS1 MARs (pGA984) (Table 1).

Each independent line was assessed by Southern analysis to characterise the integration patterns and copy number of the *gusA* and *aphIV* genes (Fig. 1). Genomic DNA from the plant lines without MARs was digested using flanking restriction enzymes expected to release a transgene fragment of discrete size (Vain et al. 1999). The first probing was performed using the R2272 RFLP probe in order to normalise genomic DNA loading. The second probe used corresponded to the *gusA* gene (Fig. 1A). More than 80% of the lines exhibited complex banding patterns in addition to the expected intact 35S :: *gusA* :: nosT unit (overall 5.72 ± 1.25 bands per transformed line, $P < 0.05$). Many lines exhibited large variation in the relative intensity of bands composing their hybridization pattern. Such complex integration patterns are commonly observed after direct DNA transfer and reflect complex transgenic loci containing multiple intact and/or modified copies of the transgene as well as interspersed genomic DNA (Lyznik et al. 1989; Gordon-Kamm et al. 1990; Svitashev and Somers 2001). However banding patterns should be interpreted with caution as they rarely allow complete and true reconstitution of the transgenic locus configuration generated by particle bombardment. The CaMV35S promoter has been described as a possible hot spot for recombination in transgenic rice plants (Kohli et al. 1999). In our study, comparison between *Eco*RI/*Hin*dIII digests (releasing the entire 35S :: *gusA* :: nos expression unit) and *Eco*RI/*Xba*I digests (releasing the *gusA* :: nos sequence) suggested that 41% of the lines showed rearrangements compatible with recombination in the CaMV35S promoter region leading to head to head expression units (Fig. 1A). The third probe used corresponded to the *aphIV* gene (Fig. 1B). Most lines contained an intact 1.3-kbp fragment corresponding to the *aphIV* gene but also additional bands of higher and lower molecular weight (4.77 ± 1.24 bands per transformed line, $P < 0.05$).

The presence of flanking MARs appeared to generate less complex banding patterns for the *gusA* gene. However, no significant difference could be found in the number of hybridizing bands in the presence or in the absence of MARs, for either the *gusA* gene (flanked by MARs) or the *aphIV* gene (unflanked by MARs) (Kruskal-Wallis and Mann-Whitney tests on the number of bands in the presence vs the absence of MARs, $P < 0.05$).

Across the entire set of 83 independent lines, with or without MARs, the integration patterns of the *gusA* gene (4.49 ± 0.64 bands per transformed line, $P < 0.05$) were significantly more complex than those of the *aphIV* gene (3.55 ± 0.58 bands per transformed line, $P < 0.05$) (Mann-Whitney test, $P < 0.05$). However, there was no significant correlation between the complexity of the integration pattern (i.e. the number of bands) of the *aphIV* and the *gusA* genes across the independent transgenic lines (Fig. 2A).

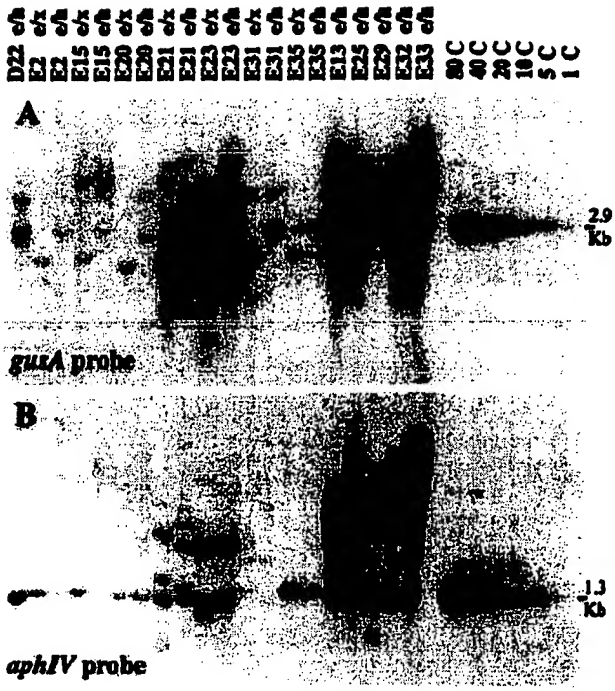


Fig. 1A, B Southern-blot analysis of transformed rice plants (T_0). Southern blots of independent transgenic plant lines co-transformed with pGHNC12 and PJIC201. Membrane **A** was probed with the *gusA* gene and membrane **B** was probed with the *aphIV* gene. Plant genomic DNA and reconstitution standards were digested using a combination of flanking restriction enzymes [*Eco*RI + *Hin*dIII (e/h) or *Eco*RI + *Xba*I (e/x)] to release a discrete size fragment from the *gusA* or *aphIV* expression unit (Vain et al. 1999). Reconstitution standards were prepared by serial dilution of pGHNC12 and pJIC201 plasmids into wild-type genomic DNA so as to introduce 1 to 80 *gusA* and *aphIV* gene copies per 2C equivalent. Expected discrete fragment sizes are indicated

Transgene copy number in T_0 rice plants

The *gusA* and *aphIV* gene copy numbers were determined for each transformed line by Southern analysis followed by densitometry. For most lines, genomic DNA of T_0 plants was analysed using two different restriction digests (Fig. 1). The error in copy number evaluation was estimated by the variation in a *gusA* positive control line present on each blot (10 ± 2.4 copies; $P < 0.05$, Vain et al. 1999). Previous studies showed that gene copy number-determination by Southern analysis plus densitometry and by quantitative PCR techniques were convergent and reliable in tobacco plants transformed with the pGHNC12 gene construct (Ülker et al. 1999). In the present study, the number of bands was not used to calculate the transgene copy number due to lack of reliability.

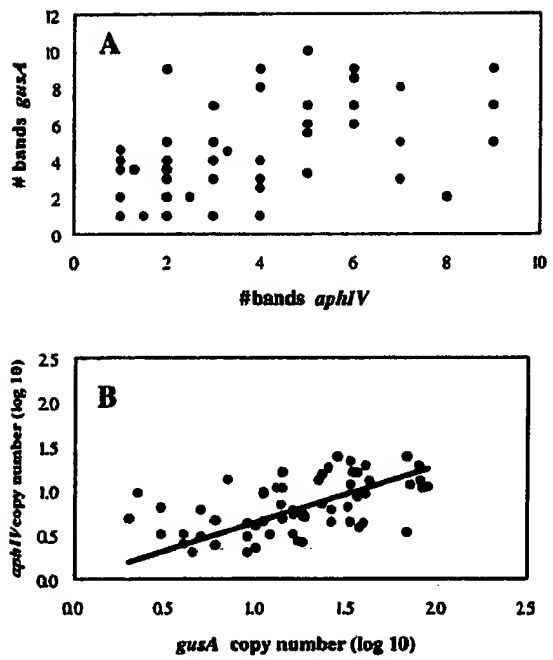


Fig. 2A, B Relationship between *gusA* and *aphIV* gene hybridisation pattern and copy number in T_0 rice plant lines. **A** Banding patterns were obtained from Southern analysis as described in Fig. 1 using restriction enzymes that are expected to release a transgene fragment of discrete size. **B** Linear regression analysis of *gusA* and *aphIV* gene copy numbers was performed using all lines with or without MARs containing both transgenes. Data were log10 transformed to ensure normality of distributions. Each data point represents one independent plant line

The *gusA* and *aphIV* gene copy numbers are shown in Table 1. In the population without MARs, the transformed plant lines contained on average 34 *gusA* and 6 *aphIV* gene copies. Forty one percent of these lines contained 1 to 20 *gusA* copies, 32% contained 20 to 40 *gusA* copies and 27% contained more than 40 *gusA* copies; 76% of the lines contained 1 to 10 *aphIV* copies and 24% of the lines contained 10 to 20 *aphIV* copies.

As previously reported (Vain et al. 1999), the presence of flanking MARs tended to reduce average *gusA* gene copy number (Table 1). However, it is difficult to draw conclusions from these average figures as very often the further relationship between copy number and other factors, such as transgene expression level, is not uniform across the entire population of transformants (Ülker et al. 1999; Vain et al. 1999). As expected, the *aphIV* gene copy number was not significantly different in the presence or absence of MARs (Table 1).

Across the entire set of 83 independent lines with or without MARs, there was a significant correlation (r = 0.33, $P < 0.01$) between the *aphIV* and *gusA* gene copy numbers (Fig. 2B). The rice plants also contained significantly more *gusA* gene copies (average of 25) than *aphIV* gene copies (average of 8) (Mann-Whitney test, $P < 0.05$). Interestingly, this ratio corresponded to the 3:1 molar ratio of *gusA*- and *aphIV*-containing plasmids used in the transformation experiments. This suggests that the number of gene copies integrated into the plant genome after particle bombardment may be directly proportional to the quantity of DNA used during the transformation process. This may provide a means of decreasing transgene copy number in transformation experiments by decreasing the quantity of DNA delivered during each shot.

Transgene expression in T_0 rice plants

Expression levels in transformed rice plants from each independent line were characterised by fluorometric GUS assay. To minimise the influence of environmental conditions, five different T_0 plants regenerated from each independent line were grown in a fully controlled growth room and analysed at the same developmental stage (five-leaf stage). Standardization of quantification of transgene expression was of particular importance as expression levels can vary during the plant life cycle and between leaves/tillers at different developmental stages (data not shown).

The average GUS expression level in the lines without MARs was 1,158 pmol MU min^{-1} mg^{-1} of protein, and varied from 0 to about 5,000 pmol MU min^{-1} mg^{-1} of protein (Fig. 3, Table 1). Despite the presence of the *gusA* transgene in all these plant lines, 41% of the transformed lines (12/29) did not exhibit GUS activity. The absence of transgene expression in primary transgenic plants is often the main factor impairing overall transgene expression in populations of plant lines generated by particle bombardment (Gordon-Kamm et al. 1990; Register et al. 1994; Vain et al. 1999). The absence of transgene expression can result from a combination of structural and epigenetic mechanisms (Finnegan and McElroy 1994; Meyer 1995; Matzke and Matzke 1998; Iyer et al. 2000). Transgene expression levels in the remaining population of expressing lines (17/29) did not exhibit obvious distribution discontinuity (Fig. 3). There was large variation in transgene expression levels between independent plant lines (coefficient of variation inter-expressing lines = 73%, Table 1). The overall expression profile of the population of independent plant lines without MARs was comparable to those previously published in other transgenic studies (Jones et al. 1985; Gordon-Kamm et al. 1990; Peach and Velten 1991). The variability of GUS activity among the five T_0 plants regenerated from each line was strongly dependent upon the mean (significant correlation of intra-line variance and mean expression level r = 0.82, $P < 0.001$) and therefore was assessed using the coefficient of variation (CV). The average CV among T_0 plants regenerated from the same line was 50% ± 15% ($P < 0.05$, Fig. 3). This variability could mostly be attributed to inter-experiment variations of GUS assay measurements, as similar CV values (40% to 60%) were obtained among 43 wild-type and 18 positive control plants, respectively, across different experiments.

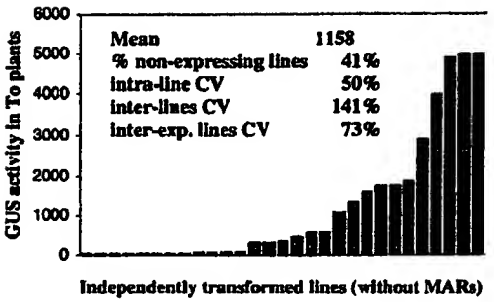


Fig. 3 Distribution of transgene expression levels in T_0 transgenic rice plant lines. Each data point represents the average *gusA* expression level of five T_0 rice plants from each independent line transformed with the pGHNC12 (35S :: *gusA* :: NOS) construct. GUS activity is expressed in pmol MU min^{-1} mg^{-1} of protein. *CV* = coefficient of variation (standard deviation/mean). *Exp* = expressing the *gusA* gene

In the population of plant lines without MARs, the *aphIV* and *gusA* genes appeared to be expressed independently since all lines exhibited strong hygromycin resistance but only 59% of the lines expressed the *gusA* gene. Independent expression of co-transformed transgenes is common to many transformed plants (Gordon-Kamm et al. 1990; Mlynárová et al. 1995; Vain et al. 1998). There was no significant difference in *aphIV* copy number or fertility between lines expressing and not expressing the *gusA* transgene ($P > 0.05$, ANOVA). However, non-expressing lines contained significantly more *gusA* gene copies (52 copies) than expressing lines (23 copies) ($P = 0.016$, ANOVA). This suggests that high gene copy number could affect transgene expression levels.

As previously reported (Vain et al. 1999), the presence of flanking MARs had a pronounced effect on transgene expression in T_0 rice plants and significantly altered the expression profile of the population of transgenic plant lines. Flanking MARs significantly reduced the occurrence of non-expressing lines from 41% (12/29) in the absence of MARs, to 17% (5/29) and 28% (7/25) for Rb7 and ARS1 MARs respectively (Table 1). Flanking MARs also significantly increased average GUS activity by 2.8-fold in the overall populations and by 2.1-fold among expressing lines (n and ne in Table 1). Flanking MARs also increased the maximal level of *gusA* expression by up to three-fold. Variation in GUS expression levels between independent lines was similar both in the presence and in the absence of MARs (CVs not significantly different, Table 1). This absence of normalisation of transgene expression at the population level is not surprising. As hypothesised by the loop model (Mirkovitch et al. 1984), flanking MARs should reduce the variability of transgene expression between independent transformants carrying the same number of active transgene copies. The expression levels of transgenes flanked by MARs should also vary in direct proportion to the active copy number. Such MAR effects should not necessarily lead to a decrease in the variability of transgene expression in the entire population of transgenic lines containing MARs. When transgenic plants are produced by direct transfer of DNA, the range of copy numbers is so broad that the copy number dependence of transgene expression provided by flanking MARs can increase the maximal expression level and is therefore not expected to (and generally doesn't) decrease the variability of transgene expression in the entire population of transgenics (Allen et al. 2000). Only when the range of transgene copy numbers is limited (e.g. after *Agrobacterium*-mediated transformation) can the copy number-dependent and position-independent transgene expression produced by MARs lead to an overall decrease of transgene expression variability in the entire population of transgenic plants (Mlynárová et al. 1994).

Among the entire set of 83 independent plant lines, with or without MARs, there was no significant difference in *aphIV* copy number or fertility between lines expressing and not expressing the *gusA* transgene ($P = 0.246$ and $P = 0.29$, respectively, ANOVA). However, non-expressing transgenic plant lines contained significantly ($P = 0.003$, ANOVA) higher *gusA* copy numbers (39 copies) than expressing ones (20 copies).

Transgene expression vs copy number in T_0 rice plants

Comprehensive analysis of the relationship between *gusA* gene copy number and expression levels was carried out for each independently transformed plant line (Vain et al. 1999). In the population of lines without MARs, despite the fact that non-expressing lines contained significantly more *gusA* gene copies than expressing lines (see previous section), there was no correlation between expression level and *gusA* gene copy number across the entire population of lines without MARs ($r = 0.3$, $P > 0.05$). This is comparable to previous studies of transgenic cereals transformed by direct transfer of DNA (Linn et al. 1990). Flanking Rb7 MARs provided copy number-dependent expression of the *gusA* transgene (up to 20 copies), but expression was generally reduced in lines carrying a higher copy number (Vain et al. 1999) probably due to silencing phenomena or transcriptional limitations (Mlynárová et al. 1995; Allen et al. 2000). This is in agreement with parallel studies using flanking Rb7 MARs in tobacco plants (Ülker et al. 1999). Overall Rb7 MARs exhibited some, but not all, of the characteristics predicted by the loop model (Mirkovitch et al. 1984). In contrast, the ARS1 MARs had a more limited "MAR effect" by exhibiting a less significant copy number-dependence of transgene expression (Vain et al. 1999).

Transgene inheritance at the structural level

T_1 seeds were obtained by self-pollination of primary transformed rice (T_0) plants. Only 61% of the initial co-transformed plant lines (51 out of 83 lines) were sufficiently fertile (i.e. more than 40 seeds) to allow further inheritance analysis (nf in Table 1).

Transgene inheritance at the structural level was analysed in up to 40 T_1 plants from each of 20 lines and in 5 to 8 T_1 plants from each of the remaining 31 fertile lines. Seeds were germinated without hygromycin selection and tested for the presence of the *gusA* and *aphIV* genes by PCR analysis (data not shown). Around 65% of the fertile lines analysed exhibited Mendelian inheritance of the *gusA* and *aphIV* transgenes at the structural level. Segregation frequencies indicated that transgenes were generally linked and integrated at only one locus (3:1 ratio after self-pollination). Two lines showed transgene integration in at least two unlinked loci (15:1 ratio after self-pollination) each containing both *gusA* and *aphIV* transgenes. When transmitted to the offspring, the *gusA* and *aphIV* transgenes always co-segregated in the T_1 progenies. Skewed segregation at the structural level was observed in around 35% (17/51) of the fertile lines tested. The progenies from these transformation events were analysed using PCR primers aimed at amplifying different regions of the *gusA* or *aphIV* transgenes. Detailed structural analysis showed that many plant progeny did not contain, or only contained fragments, of the *gusA* and/or *aphIV* transgenes. Among fertile lines, 18% (9/51) showed no transmission of *gusA* nor *aphIV* transgenes to any of their progeny, despite repeated confirmation that both intact transgenes were present in the parent T_0 transgenic plants by Southern analysis (Fig. 1) and by expression studies (Table 1). Six percent (3/51) of the lines (such as D28 described in the next section) exhibited transgene inheritance at a significantly lower frequency than 3:1 (as expected after self-pollination) but transgenes were co-inherited without detectable alteration of expression units. Ten percent (5/51) of the lines exhibited complex segregation patterns with some T_1 plants containing no transgene and others containing only fragments of the original *gusA* and/or *aphIV* transgenes. The occurrence of lines with this latter type of behaviour suggested that some transgenic loci generated by particle bombardment are likely to be altered from one generation to the next through recombination or deletion. Such processes could also have occurred without being detected in high-copy number-lines exhibiting Mendelian inheritance, as long as at least one intact expression unit of each transgene was inherited. Interestingly, the fertile lines showing non-transmission of intact transgenes to some or all progenies (17/51) exhibited significantly lower fertility (51 vs 124 seeds, $P = 0.007$, ANOVA) but the same *gusA* and *aphIV* gene copy number (23 vs 26 *gusA* gene copies and 10 vs 7 *aphIV* gene copies, $P > 0.05$, ANOVA) and the same *gusA* expression level (4,651 vs 3,703 pmol MU min^{-1} mg^{-1} of protein, $P = 0.176$, ANOVA) as lines transmitting intact transgenes. Similar loss of transgenes from one generation to the next has also been reported in transformed maize plants (Walters et al. 1992) and associated with low plant fertility (Register et al. 1994).

Non-transmission of intact transgenes to some or all progenies affected at least 14% of the lines in the population of plants without MARs. The presence of flanking MARs had no significant effect on this situation (21% and 28% were affected by poor transgene transmission in the presence of Rb7 and ARS1 MARs respectively). Non-transmission of intact transgenes to some or all progenies was a key factor in generational transgene instability. When it occurs at a low frequency it can easily been confused with levels of gene silencing if no structural study is conducted in parallel to the segregation study at the expression level. It can also be mistaken for out-segregation of co-transformed transgenes unless multi-primer PCR or Southern analysis is conducted on progenies. Study of transgene stability at the structural level was a prerequisite for characterising other forms of instability at the expression level, such as reduced expression levels or gene silencing.

Generational stability of transgene expression

Stability of transgene expression was first assessed by segregation analysis, then by quantitative measurement of transgene expression levels in progenies. Only 51 of the initial 83 co-transformed plant lines exhibited sufficient levels of fertility (i.e. more than 40 seeds) and only 34 of these fertile lines transmitted both transgenes to their progenies, allowing monitoring of *aphIV* and *gusA* gene expression across generations. Twenty one of the 34 lines expressed the *aphIV* and *gusA* genes in T_0 plants allowing inheritance studies of transgene expression to be conducted in parallel for both transgenes (Fig. 4A). The remaining 13 lines expressed the *aphIV* gene but not the *gusA* gene in T_0 plants, and therefore a transgene inheritance study at the expression level could only be conducted for the *aphIV* gene (Fig. 4B). More than 5,100 segregating T_1 seedlings were tested for expression of both transgenes in the T_1 generation. Segregation frequencies indicated that 28 (82%) and 17 (81%) lines showed Mendelian inheritance (1 or 2 loci, $P < 0.05$, chi-square analysis) at the expression level for the *aphIV* and the *gusA* genes, respectively (Fig. 4). Most of the remaining lines showed skewed inheritance at the expression level for only one of the transgenes. Interestingly, significantly more non-expressing GUS lines (4/13) exhibited skewed segregation for hygromycin resistance than did expressing lines (2/21) ($P < 0.05$, chi square analysis, Fig. 4A vs 4B). This suggests that inactivation of one transgene (*gusA*) could be associated with the destabilisation of expression of the other transgene (*aphIV*) co-integrated at the same locus (Fig. 4B). Segregation frequencies could not be correlated with the variability of GUS activity (CV) among T_0 plants regenerated from each line ($r = 0.16$, $P > 0.05$). However, lines with skewed segregation exhibited significantly higher intra-line *gusA* expression variability in parent T_0 plants than lines with Mendelian inheritance (Mann-Whitney skewed vs non-skewed segregation, $P = 0.014$). This suggests that variability of transgene expression among clonal T_0 plants may be a sign of further generational transgene instability.

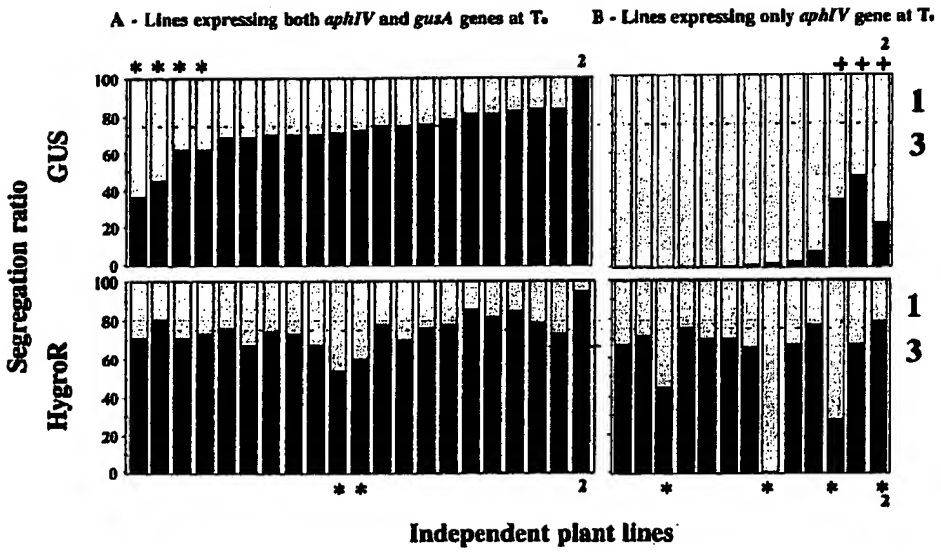


Fig. 4A, B Segregation analysis of transgene expression in T_1 transgenic rice seedlings. Each bar represents the segregation ratio of at least 40 seedlings expressing (in *black* or *dark grey*) or not expressing (in *light grey*) the transgenes. Among the 51 fertile plant lines with or without MARs studied, 17 lines did not transmit intact transgenes to their progeny (data not represented here), 21 lines expressed both the *gusA* and *aphIV* gene in the T_0 plants (A) and 13 expressed the *aphIV* but not the *gusA* gene in T_0 plants (B). *Star* (*) indicates the skewed segregation ratios significantly different from Mendelian 3:1 or 15:1 ratios (chi-square test, $P < 0.05$). *Cross* (+) indicates lines not expressing the *gusA* gene in T_0 and T_1 plants but expressing the *gusA* gene in vitro and in T_1 seeds. Two (2) indicates lines containing two independent transgenic loci

Among the 51 fertile lines with or without MARs, almost half of the lines showed skewed segregation at the expression level for the *aphIV* gene (45% of the lines) or the *gusA* gene (50% of the lines). This is comparable to other inheritance studies measured by the transgene phenotype in cereals (Register et al. 1994; Pawlowski and Somers 1996). In this study, deviation from Mendelian segregation of transgenes at the expression level was due to a combination of poor transmission of the transgene(s) (see previous section) and/or transgene silencing. In line D28, among 41 T_1 plants tested for the presence and expression of the *aphIV* and *gusA* genes, 44% (19/41) contained and expressed the transgenes, 10% (4/41) contained but did not express the transgenes (silencing) and 46% (18/41) did not contain the transgenes (25% expected to be non-transformed segregants + 21% non-transmission of transgenes). In lines such as D28 the non-transmission of transgenes (21% of T_1 plants) had a greater impact than transgene silencing (10% of T_1 plants) on skewing segregation ratios determined at the expression level. Skewed segregation of transgene expression was observed in the presence or in the absence of MARs suggesting that MARs do not eliminate skewed inheritance or transgene silencing.

Stability of transgene expression level was then assessed by comparison of GUS activity in T_0 and T_1 plants from the 21 lines containing and expressing the transgenes at both generations. Lines exhibiting Mendelian segregation at the structural level and skewed segregation at the expression level systematically showed reduced or silenced transgene expression levels in some T_1 plants. In the absence of MARs (pGHNC12), *gusA* expression levels were significantly heritable ($P < 0.001$); however, there was a significant 35% reduction of transgene expression levels in T_1 plants compared to the parental T_0 plants (Vain et al. 1999). Three lines exhibited an unstable transgene expression level in the progeny due to gene silencing. The overall impact of gene silencing in the entire population of plant lines without MARs (29 lines) was therefore around 10%.

As previously reported (Vain et al. 1999), the presence of flanking Rb7 MARs appeared to significantly improve the stability of transgene expression levels over two generations at the population level. These observations are similar to those seen in parallel studies in tobacco plants using the same experimental procedure (same gene construct introduced by particle bombardment, Ülker et al. 1999). Nevertheless in rice, generational instability of transgene expression was not eliminated by the presence of flanking MARs. In the population of plant lines containing MARs, 13% of the lines (7/54) exhibited an unstable (e.g. reduced or silenced) transgene expression level in the progeny. However, among these lines, only 7% (4/54) showed transgene silencing in some or all progenies. The remaining 6% (3/54) of lines showed only a decrease in transgene expression without any progeny being silenced. The overall benefit of flanking MARs in a population of transgenic rice plants was therefore mostly to increase the occurrence of stable lines (from 7% to 17%) rather than to reduce the production of unstable lines (from 10% silenced to 7% silenced plus 6% reduced expression). In other plant species, MARs have been shown not to protect

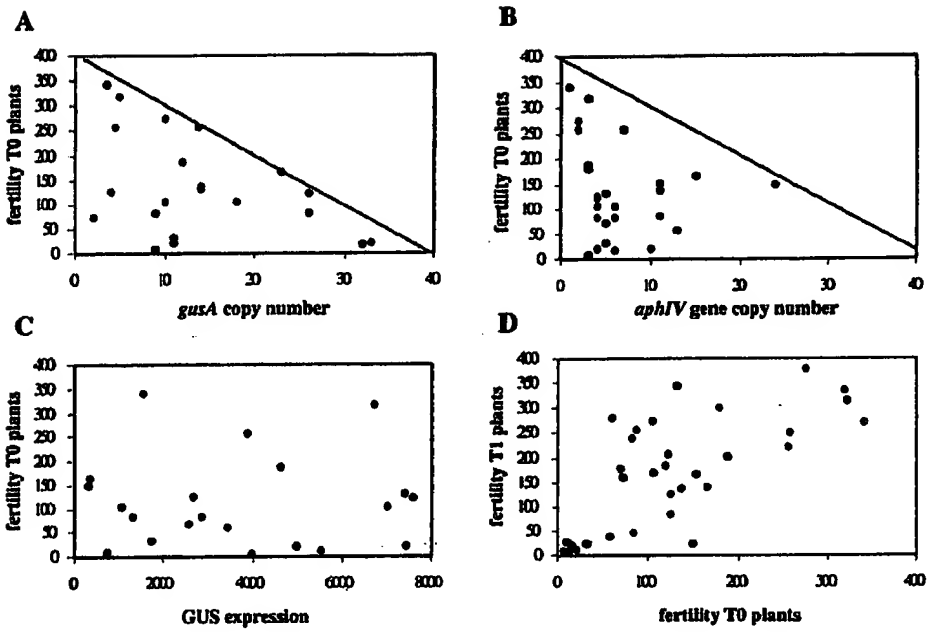


Fig. 5A–D Factors influencing T_0 and T_1 rice plant fertility. Fertility of each independently transformed plant line was determined by the average number of seeds produced by five T_0 and eight T_1 rice plants from each line. A and B Relationship between transgene copy number and T_0 rice plant fertility. C Relationship between transgene expression level and T_0 rice plant line fertility. D Fertility of T_0 vs T_1 rice plants. Each data point represents one independent plant line

against strong post-transcriptional gene silencing (Allen et al. 2000).

In the 21 lines, with or without MARs, expressing both the *aphIV* and *gusA* genes in T_0 and T_1 plants, a decrease/silencing of transgene expression across generations seemed independent of *gusA* expression level, *aphIV* copy number and fertility (Kruskal-Wallis and ANOVA tests on stable vs unstable lines, $P > 0.05$). However transgenic lines with unstable GUS expression contained significantly higher *gusA* copies than lines with stable GUS expression (average of 23 and 11 *gusA* copies respectively, Kruskal-Wallis and ANOVA, $P = 0.039$). This suggests that high copy number might interfere with transgene stability over generations. The importance of transgene copy number on transgene expression stability was confirmed when lines exhibiting other forms of instability such as those at the structural level (e.g. non-transmission of intact transgenes to some or all progenies) were also included in the analysis. Lines showing instability at the structural or expression levels contained significantly more *gusA* and *aphIV* gene copies (Kruskall-Wallis unstable vs stable lines $P = 0.032$ and $P = 0.017$ respectively).

Fertility of transformed rice plants

The fertility of transformed rice plants from each independent line was measured over two generations. In the first generation (T_0 plants), 55% (16/29) of independently transformed lines without MARs consistently produced fertile plants (84 seeds per T_0 plant, Table 1). The remaining lines (13/29) were sterile. In this population, 28% (8/29) of the lines were sterile expressers. Across the entire set of 83 independent lines with or without MARs, 61% (51/83) of the lines consistently produced fertile plants (nf in Table 1). There was no significant difference in fertility levels between the expressing (89 seeds per plant line on average) and non-expressing lines (120 seeds per plant line on average) ($P = 0.225$, ANOVA). Only 34 of the 51 fertile lines exhibited inheritance of both transgenes allowing monitoring of plant fertility in transformed plants across generations (nfh in Table 1). The presence of flanking MARs did not affect the fertility of transgenic plants ($P = 0.553$, ANOVA). Among lines, with or without MARs, there was no significant difference in GUS activity between sterile and fertile plants ($P = 0.132$, ANOVA). There was no correlation ($r = 0.009$, $P > 0.1$, Fig. 5C) between the fertility level of the transgenic plants and the expression level of the *gusA* gene. There was also no significant difference between fertile and sterile plants for *gusA* and *aphIV* copy number ($P > 0.05$, ANOVA). However, high *aphIV* or *gusA* gene copy numbers were never associated with high fertility (Fig. 5A and B). This suggests that high transgene copy number could interfere with gamete or seed development/viability. Since each primary transformant is hemizygous for the transgene(s), it is possible that long stretches of foreign DNA with no homology to the homoeologous chromosome might interfere with meiosis. In the second generation, the fertility of T_1 plants (161 seeds per plant line on average) was significantly improved (by 30%) when compared to the fertility of the T_0 plants ($P = 0.01$, *t*-test of means on fertility T_0 to fertility T_1 across independent lines). Improvement in fertility level in the progeny of primary transgenics has often been reported in studies on transgenic cereals, but this was, however, slight in the present study. Among the

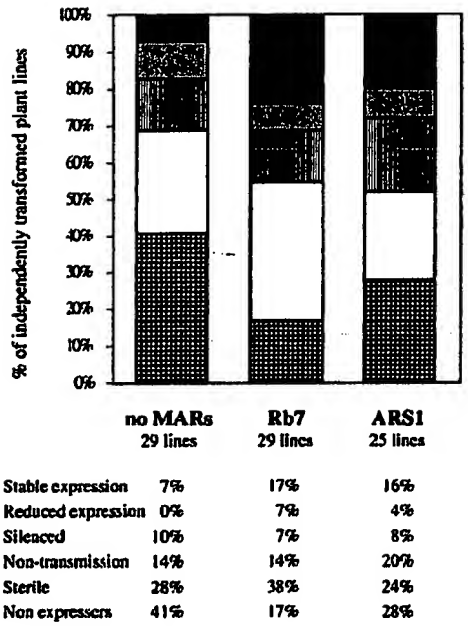


Fig. 6 Generational stability of transgene expression in rice plant lines, in the presence or in the absence of flanking MARs (from T_0 to the T_1 generation). T_0 plants were all hemizygous. T_1 plants were expected to be 2/3 hemizygous and 1/3 homozygous. Non expresser: line not expressing the *gusA* gene in all T_0 plants. Non-transmission: non-transmission of intact transgenes to some or all T_1 plants. Silenced: line expressing the *gusA* gene in all T_0 plants but with at least one T_1 plant silenced (containing but not expressing the transgenes). Reduced expression: no silencing observed in any T_1 plant

34 independent fertile lines transmitting the transgene(s) with or without MARs, no significant correlation could be shown between fertility levels at the T_0 and T_1 generations as the data set was not normally distributed even after data transformation. However, lines producing plants with low fertility levels consistently exhibited low fertility across generations (Fig. 5D).

Conclusion

Transgene behaviour in the population of transgenic rice plants generated in this study has been summarised in Fig. 6. Over two generations, only a small proportion of the plant lines without MARs (7%) exhibited Mendelian inheritance and stable expression of the unselected transgene (*gusA*). Transgene inactivation occurred in primary transgenic plants in 41% of the lines. In the next generation (T_1), loss or reduction of transgene expression in plants was mostly due to plant sterility (28% of lines), non-transmission of intact transgenes to some or all progenies (14% of lines) and transgene silencing (10% of lines). This profile is in accordance with many aspects of transgene behaviour described in other transgenic plant studies using particle gun bombardment (Pawlowski and Somers 1996; Ülker et al. 1999). The occurrence of transgene instability could also increase in subsequent generations (Kumpalta and Hall 1998) especially when plants are stabilised at the homozygous level (James et al. 2002). Alternative experimental conditions, transgene constructs, expression assays or plant material can negatively or positively influence transgene behaviour and consequently modulate this profile. Nevertheless, to-date most transgenic plants generated by direct transfer of DNA exhibit similar high levels of transgene instability, which probably originates from their common type of complex transgenic locus structure. Molecular studies have shown that direct DNA transfer often leads to integration at one locus (rarely two loci) of multiple fragmented and rearranged transgene copies as well as plasmid backbone sequences (Gordon-Kamm et al. 1990; Wan and Lemaux 1994). Massive rearrangements of genomic DNA including large scale duplication, deletion and translocation were also observed at the integration site (Takano et al. 1997). Extensive scrambling of transgene and intervening genomic DNA sequences has been identified by fiber-FISH experiments in oat plants transformed by particle bombardment (Svitashev and Somers 2001). The presence of plasmid backbone sequences in transgenic loci probably has a strong effect on transgene expression level and stability (Fu et al. 2000). However, it remains unclear whether the removal of plasmid backbone is enough to create simple transgenic loci through direct transfer of DNA (Breitler et al. 2002).

In this study, aspects of plant development and transgene stability across generations were significantly influenced by the copy number and to a further extent by the expression level of the unselected marker gene (*gusA*) itself. High *gusA* copy number significantly decreased transgene expression level and the stability of transgene expression across generations, and to a lesser extent plant fertility. The absence of *gusA* expression was associated with silencing of the co-transformed hygromycin resistance gene across generations. Sequential analysis of transgene inheritance at the structural then the expression levels, followed by quantification of transgene expression level in progeny, was crucial in identifying different types of transgene behaviour and characteristics. In this study, transgene instability at the structural level (i.e. non-transmission of intact transgenes to some or all progenies) was as important, if not more important, than instability at the expression level (reduced expression, silencing). Transgene inactivation, was also clearly a dynamic process occurring at any step of plant development or generation (data not shown). Transformation events exhibiting intra-line or developmental transgene instability in primary T_0 plants, were often affected by subsequent generational transgene instability. The structural and epigenetic mechanisms underlying such transgene instability in plants have been extensively described (references in Introduction) but to-date, the relative contribution of these mechanisms to transgene inactivation still remains unclear at the population level.

In this study, MARs exhibited some, but not all, of the characteristics predicted by the loop model. Flanking Rb7

MARs showed copy number-dependent transgene expression up to 20 gene copies and some reduction in position effects. In addition, and not directly predicted by the loop model, MARs increased the overall occurrence of stable lines over two generations. Studies over more generations, at different ploidy levels and using larger populations of transformed lines, will determine if this is a consistent feature of flanking MARs in transgenic studies.

To-date, the control of transgene integration, structure, and subsequent expression levels and stability remain key issues limiting transgenic studies. In cereals, the production of single-copy transgenic loci at high frequencies via *Agrobacterium*-mediated transformation (Hiei et al. 1994) and the prediction of transgene expression levels through copy number dependence using flanking MARs (Vain et al. 1999) have made a contribution towards this goal. They will also contribute to further improvements of cereal transgenic studies through unlinked transgene integration (Komari et al. 1996), the resolution of transgenic loci by secondary modifications (Srivastava et al. 1999) or by transposon-based strategies and gene targeting (Pazkowski et al. 1988).

Acknowledgements We thank WARDA (Ivory Coast) for providing ITA212 rice seeds. We gratefully acknowledge The John Innes Foundation for its support and A. Yang for technical assistance. This document is an output from a project (Plant Sciences Research Programme R6948) funded by the UK Department for International Development (DFID) and administered by the Centre for Arid Zone Studies (CAZS) for the benefit of developing countries. The views expressed are not necessarily those of DFID.

References

Allen GC, Spiker S, Thompson WF (2000) Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Mol Biol 43:361–376

Bhattacharyya M, Stemer BA, Dixon RA (1994) Reduced variation in transgene expression from a binary vector with selectable markers at the right and left T-DNA borders. Plant J 6:957–968

Breitler JC, Cordero MJ, Royer M, Meynard D, Segundo BS, Guideroni E (2002) The -689/+197 region of the maize protease inhibitor gene directs high level, wound-inducible expression of the cry1B gene which protects transgenic rice plants from stemborer attack. Theor Appl Genet (in press)

Breyne P, Gheysen G, Jacobs A, Van Montagu M, Depicker A (1992) Effect of T-DNA configuration on transgene expression. Mol Gen Genet 235:389–396

Christou P (1996) Transformation technology. Trends Plant Sci 1:423–431

Finnegan J, McElroy D (1994) Transgene inactivation: plants fight back. Bio/Technology 12:883–888

Fu X, Duc LT, Fontana S, Bong BB, Tinjuangjun P, Sudhakar D, Twyman RM, Christou P, Kohli A (2000) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. Transgenic Res 9:11–19

Gordon-Kamm WJ, Spencer TM, Mangano ML, Adams TR, Daines RJ, Start WG, O'Brien JV, Chambers SA, Adams WR Jr, Willetts NG, Rice TB, Mackey CJ, Krueger RW, Kausch AP, Lemaux PG (1990) Transformation of maize cells and regeneration of fertile transgenic plants. Plant Cell 2:603–618

Hansen G, Chilton M-D (1996) "Agrolistic" transformation of plant cells: Integration of T-strands generated in planta. Proc Natl Acad Sci USA 93:14,978–14,983

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries T-DNA. Plant J 6:271–282

Jefferson RA, Kavanagh TA, Bevan MW (1987) β-glucuronidase as a sensitive and versatile fusion marker in higher plants. EMBO J 6:3901–3907

Iyer LM, Kumpatla SP, Chandrasekharan MB, Hall TC (2000) Transgene silencing in monocots. Plant Mol Biol 43:323–346

James VA, Avart C, Worland B, Snape JW, Vain P (2002) The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. Theor Appl Genet 104:553–561

Jones JDG, Dunsmuir D, Bedbrook J (1985) High level expression of introduced chimearic genes in regenerated transformed plants. EMBO J 10:2411–2418

Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. Plant J 17:591–601

Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. Plant J 10:165–174

Kumpatla SP, Hall TC (1998) Recurrent onset of epigenetic silencing in rice harboring a multi-copy transgene. Plant J 14: 129–135

Linn F, Heidmann I, Saedler H, Meyer P (1990) Epigenetic changes in the expression of the maize A1 gene in *Petunia hybrida*: Role of numbers of integrated gene copies and methylation state. Mol Gen Genet 222:329–336

Lyznik LA, Ryan RD, Ritchie SW, Hodges TK (1989) Stable co-transformation of maize protoplasts with *gusA* and *neo* genes. Plant Mol Biol 13:151–161

Matzke AJM, Matzke MA (1998) Position effect and epigenetic silencing of plant transgenes. Curr Opin Plant Biol 1:142–148

Meyer P (1995) Understanding and controlling transgene expression. Trends Biotechnol 13:332–337

Mirkovitch J, Mirault ME, Laemmli UK (1984) Organization of the higher-order chromatin loop: DNA attachment sites on nuclear scaffold. Cell 39:223–232

Mlynárová L, Loonen A, Heldens J, Jansen RC, Keizer P, Stiekema WJ, Nap JP (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell 6:417–426

Mlynárová L, Jansen RC, Conner AJ, Stiekema WJ, Nap JP (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. Plant Cell 7:599–609

Pazkowski J, Baur M, Bogucki A, Potrykus I (1988) Gene targeting in plants. EMBO J 7:4021–4026

Pawlowski WP, Somers DA (1996) Transgene inheritance in plants genetically engineered by microprojectile bombardment. Mol Biotechnol 6:17–30

Peach C, Velten J (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol 17:49–60

Register JC III, Peterson DJ, Bell PJ, Bullock WP, Evans IJ, Frame B, Greenland AJ, Higgs NS, Jepson I, Jiao S, Lewnau CJ, Sillick JM, Wilson HM (1994) Structure and function of selectable and non-selectable transgenes in maize after introduction by particle bombardment. Plant Mol Biol 25:951–961

Srivastava V, Anderson OD, Ow DW (1999) Single-copy transgenic wheat generated through the resolution of complex integration patterns. Proc Natl Acad Sci USA 96:11,117–11,121

Svitashev SK, Somers DA (2001) Genomic interspersions determine the size and complexity of transgene loci in transgenic plants produced by microprojectile bombardment. Genome 44:691–697

Takano M, Egawa H, Ikeda JE, Wakasa K (1997) The structures of integration sites in transgenic rice. Plant J 11:353–361

Ülker B, Allen GC, Thompson WF, Spiker S, Weissinger AK (1999) A tobacco MAR increases transgene expression and protects against gene silencing in the progeny of transgenic tobacco plants. Plant J 18:253–263

Vain P, Worland B, Clarke MC, Richard G, Beavis M, Liu H, Kohli A, Leech M, Snape J, Christou P, Atkinson H (1998) Expression of an engineered proteinase inhibitor (Oryzacystatin-IΔd86) for nematode resistance in transgenic rice plants. Theor Appl Genet 96:266–271

Vain P, Worland B, Kohli A, Snape JW, Christou P, Allen GC, Thompson WF (1999) Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny. Plant J 18:233–242

Wan Y, Lemaux PG (1994) Generation of large numbers of independently transformed fertile barley plants. Plant Physiol 104:37–48

Walters DA, Vetsch CS, Potts DE, Lundquist RC (1992) Transformation and inheritance of a *hygromycin phosphotransferase* gene in maize plants. Plant Mol Biol 18:189–200

Appl Microbiol Biotechnol (1996) 46:1–9

MINI-REVIEW

G. Sawers · M. Jarsch

Alternative regulation principles for the production of recombinant proteins in *Escherichia coli*

Received: 17 November 1995/Received revision: 9 February 1996/Accepted: 4 March 1996

Abstract Established expression vectors exploiting regulated promoters such as the *lac* or *tac* promoters have economic and technical limitations when used for the industrial production of recombinant proteins. Consequently, alternative expression systems are being developed that can be more readily manipulated while maintaining high yields of protein. Several suitable expression vectors have been described for use in *Escherichia coli* that are based on promoters the activity of which is under metabolic control. This article discusses the advantages and disadvantages of a cross-section of these expression systems, how they compare with established systems and how they can be applied to the industrial-scale production of recombinant proteins.

Introduction

A number of criteria must be considered when optimizing conditions for the large-scale overproduction of a recombinant protein. These cover the stability of the mRNA (Gross 1989), the efficiency of mRNA translation (Gold and Stormo 1990), the accuracy of amino acid incorporation (Santos and Tuite 1993), whether the protein is correctly folded (Hockney 1994), the formation of insoluble protein aggregates (inclusion bodies) (Hockney 1994), the susceptibility of the product to proteolysis (Gottesman 1990; Nygren et al. 1994) or the requirement for post-translational modification such as proteolytic processing or phosphorylation, and whether the product must be exported to produce an active protein (Hockney 1994; Missiakis et al. 1993). All of these criteria must be considered for each product individually. The choice of promoter used in a vector system has a major bearing on many, but not all, of these criteria. Hence, factors such as inclusion-body formation, the frequency of mistranslation events or the correct folding of a protein can be influenced according to how the activity of the promoter is regulated. Furthermore, use of a strong promoter that produces large amounts of mRNA substrate can compensate in part for mRNA instability, poor translation efficiency or an unstable product. Prudent use of a promoter, therefore, can have a fundamental impact on the quality and yield of a recombinant protein while concomitantly minimizing both development and production costs. In this review we shall describe several metabolically regulated promoters that have been successfully used in expression vectors for the large-scale production of recombinant proteins in *E. coli*. The activity of these promoters responds to fluctuations in metabolism brought about by manipulating cultivation parameters such as the concentration of phosphate or oxygen. Consequently, expression vectors that include metabolically regulated promoters provide cost-effective alternatives to established expression systems based on, for example, the *lac* or *tac* promoters.

G. Sawers
Nitrogen Fixation Laboratory, John Innes Centre,
Norwich NR4 7UH, UK

M. Jarsch (✉)
Div. Therapeutika, F & E Biotechnologie,
Boehringer Mannheim GmbH, Werk Penzberg,
Nonnenwald 2, D-82372 Penzberg, Germany.
Fax: (49) 8856 602144

Categories of recombinant protein

In essence, there are three major categories of recombinant proteins currently produced on an industrial scale, and these are listed in Table 1. The first category includes technical enzymes and proteins for food processing where it is not critical that the product is absolutely pure but high yield is an important factor (Richter 1986). Therefore, since minimal downstream processing is advantageous it is preferable that the

Table 1 Principle categories of recombinant protein

Category	Qualitative and Quantitative prerequisites	Application
1	High yield Low cost Purity is not a major concern	Technical enzymes, e.g. proteases and lipases for washing powders Proteins for food processing or supplementation, e.g. glucose oxidase Biocatalysts, e.g. glucose oxidase
2	High purity Low development effort and cost Yield less critical than for category 1	Enzymes and proteins for in vivo diagnostics, e.g. cholesterol oxidase, glucose dehydrogenase or penicillin-G acylase
3	Very high quality Consistency in production Process must be compatible with Food and Drug Administration guidelines	Human therapeutics, e.g. recombinant tissue plasminogen activator, insulin

protein is produced in high amounts, of the order of 20%–40% of the total cellular protein. For this purpose it is necessary to have an expression vector with a strong promoter. In the other two categories there is a stronger emphasis on the quality of the product, quantity being a secondary consideration (Kopetzki et al. 1994). This is particularly relevant for production of therapeutics where a consistently high-quality product is a prerequisite. Again, this can be affected by the promoter driving expression, particularly, how strong the promoter is and how its activity is controlled. The protein-synthetic capacity of a cell limits the amount, and to a certain extent the quality, of a recombinant protein that can be synthesized. Therefore, although under ideal circumstances it would be desirable to have a host cell that grows rapidly and attains very high cell densities before synthesis of the recombinant protein is switched on, this is not always practicable because the host cell must deliver sufficient ATP and metabolic intermediates to synthesize large amounts of a recombinant protein. Moreover, at very high cell densities nutrients become limiting, which can lead to misincorporation of amino acids (Santos and Tuite 1993) or premature termination of polypeptide chain elongation (Balbas and Bolivar 1990). Hence, it is necessary to be able to control promoter activity easily and efficiently so that synthesis of the recombinant protein is optimal.

Relevant features of promoters in expression vectors

There are both essential and desirable features of a promoter for use in an expression vector (Table 2). It is essential when the promoter is activated that formation of the open transcription complex is efficient and that promoter clearance is rapid. This ensures that large amounts of mRNA are synthesized. The promoter should have sufficient strength to ensure that, under optimal conditions, the ultimate product can attain levels greater than 10% of the total cellular protein, and the promoter should be regulated. The form this regulation takes, the "tightness" (i.e. the extent to which promoter activity can be prevented) of that regulation and the extent to which promoter activity can be induced (the induction ratio) are important considerations.

Transcription initiation from a promoter can be regulated either positively or negatively (Fig. 1). Positive regulation means that a specific activator protein must be present either to permit RNA polymerase to initiate transcription or to increase the frequency of transcription initiation (Gralla 1990). The activity of

Table 2 Features of a promoter desirable in an expression vector

Salient features of the promoter	Alternatives, variables or requirements
Location and stability	1. High-copy-number plasmid 2. Transcription should not interfere with plasmid replication
Strength	1. Slow initiation and rapid elongation, e.g. *lacUV5* promoter 2. Fast initiation and very rapid elongation, e.g. phage T7 promoters
Regulation	1. No promoter activity until product is desired 2. Low-level expression from promoter during growth followed by controlled activation
Activation	1. Removal of a repressor 2. Activation of a positive control factor (activator)
Mode of controlling activation	1. Temperature-shift inactivation of a repressor molecule 2. Chemical inducer added to the culture medium 3. Nutrient deprivation allowing derepression or activation of the promoter

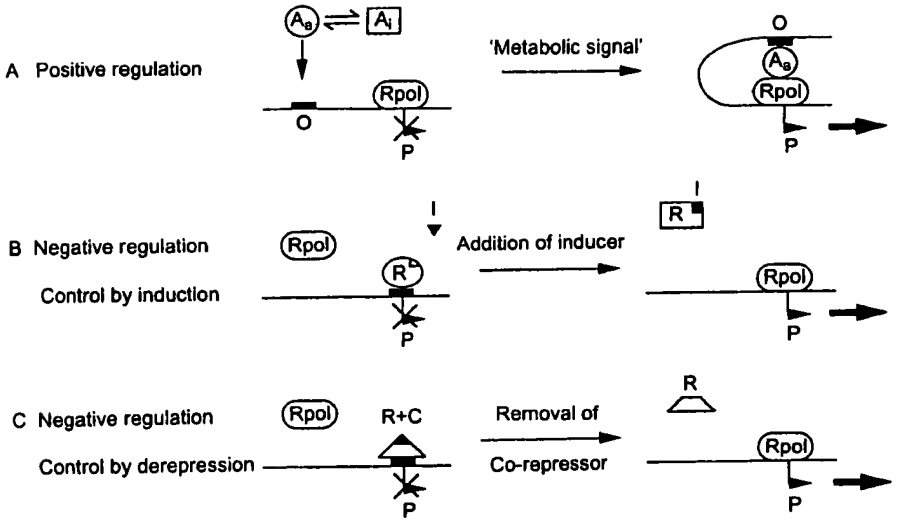


Fig. 1A–C Schematic representation of common modes of promoter regulation. **A** Positive regulation. RNA polymerase (*RPol*) can interact with the promoter (*P*) but cannot form an "open" complex (melt duplex DNA and initiate transcription) because the activator A_i is in the inactive form. Upon responding to a metabolic signal, e.g. a change in oxygen, nitrogen or phosphorus concentration, A_i undergoes a conformational change to its active conformation, A_a. A_a can then bind to a specific operator sequence (*O*; ■), which is located upstream of the promoter on the DNA. By making contact with RNA polymerase facilitated through DNA "looping" the activator can promote open complex formation and transcription initiation. Two modes of negative regulation are depicted. **B** In control by induction, the interaction of RNA polymerase with the promoter is prevented by binding of the repressor protein (*R*) to its operator site, which in this case overlaps the promoter. The repressor has a much higher affinity for its operator than RNA polymerase has for the promoter. The repressor-inducer (I) complex has a drastically reduced affinity for the operator allowing access of RNA polymerase to the promoter. **C** In the case of control by derepression, the scenario is similar to that depicted in **B** except that the repressor must interact with a small molecule (*C* = co-repressor) to enhance its affinity for its operator. Metabolism of the co-repressor inactivates the repressor and liberates the promoter.

the activator protein is usually controlled in response either to a change in the metabolic status of the cell or to the addition of a specific inducer molecule. Negative regulation means that transcription initiation from a promoter is prevented by a repressor molecule (usually a protein). The activity of a negatively regulated promoter, in turn, can be controlled in two ways. In one instance the promoter is controlled by induction, e.g. the *lac* promoter (Makoff and Oxer 1991), where the addition of the inducer lactose prevents repressor binding, thus allowing transcription to proceed. In contrast, the *trp* promoter (Squires et al. 1975), for example, is controlled by repression. Here tryptophan is a co-repressor and, when it becomes depleted, the repressor can no longer bind to the promoter and transcription can occur. How tightly a promoter is regulated depends principally to what extent transcription still occurs in the presence of a repressor or in the absence of an activator. Hence, if the recombinant protein is toxic to the host it is necessary that the promoter is very tightly regulated (Wülfing and Plückthun 1993). In the case of plasmid-based promoters controlled by repression this may require that extra copies of the gene encoding the repressor are also supplied on a plasmid to prevent repressor titration (Stark 1987).

There are several other features of a promoter that may be desirable but this will depend very much on the quality of the product, the quantity of the product that is required and whether the product is toxic to the host cell. These features are listed in Table 2.

Expression systems

Established expression systems

Strong, regulated promoters commonly employed in both research laboratories and industry to drive heterologous gene expression include the promoters from the *lac* operon and the tryptophan (*trp*) biosynthetic operon, as well as phage promoters such as the λp_L promoter and the $\phi 10$ promoter from phage T7 (Table 3). All of them have been used to produce large numbers of recombinant proteins that attain levels of at least 5%–10% of the TCP. The *trp* and *lac* promoters are both negatively regulated and have been used successfully to overproduce a vast number of recombinant proteins (Balbas and Bolivar 1990; Tacon et al. 1980; Yansura and Henner 1990).

Hybrid promoters, combining different portions of the *lac* and *trp* promoters have been constructed and used to design improved expression vectors compared with those based around the natural promoters. Examples include the *tac*, *trc* and *tic* promoters (DeBoer et al. 1983; Brosius et al. 1985) which combine the −35 RNA

Table 3 Established expression systems. *ITPG* isopropylthiogalactoside, *IAA* 3-indoleacrylic acid

Promoter	Regulation principle	Induction ratio	Advantages	Disadvantages	References
P_{trp}	Repressed by TrpR-tryptophan complex Starvation of Trp relieves repression IAA mimics Trp starvation	50- to 80-fold	Strong promoter	IAA is expensive Defined synthetic medium required if IAA is not used Potential problems with mistranslation	Squires et al. 1975; Tacon et al. 1980
P_{lac}	Repressed by LacI Induction by lactose or the gratuitous inducer IPTG Catabolite regulation	1000-fold	Strong promoter Very tight regulation IPTG is not metabolized Rapid promoter activation	IPTG is expensive Conditions must be optimized for each scale-up step in the fermentation Inclusion body formation	Gralla 1990; Squires et al. 1975; Fuller 1982
λP_L	Temperature-sensitive cI repressor	> 300-fold	Strong promoter Tight regulation	High energy expenditure Difficult to control Inclusion-body formation	Remaut et al. 1981; Rosenberg et al. 1983
Phage T7 promoters, e.g. $\phi 10$	Transcribed exclusively by phage T7 polymerase	> 1000-fold	Very strong promoter Exclusive transcription	Phage T7 RNA polymerase must be introduced into host cell and its synthesis controlled Owing to extremely high mRNA production, problems with aberrant translation can occur	Studier and Moffat 1986; Studier et al. 1990; Tabor and Richardson 1985

polymerase recognition region from the *trp* promoter with a canonical −10 RNA polymerase recognition sequence and the LacI operator from the *lac* promoter. The three promoters differ in the spacing between the −35 and −10 sequences which affects promoter strength. The *tac* promoter is the most efficient of the three, being five times stronger than *lacUV5* and it still retains the regulation by LacI (DeBoer et al. 1983; Brosius et al. 1985).

Although these expression systems are used in industry there are several shortcomings that detract from their positive attributes. As mentioned above, temperature shifts may cause localized overheating within the fermenter and they are difficult to control. Such shifts have the added disadvantage that correct protein folding is often impaired and they can exacerbate the risk of proteolysis (Hockney 1994). Inclusion-body formation also can be increased by temperature shifts; however, this is not always undesirable and depends very much on the product. Owing to the increased number of operator sites in plasmid-based expression systems it is often necessary to increase the number of repressor molecules. This can be achieved by introducing the gene encoding the repressor onto the expression plasmid itself or by using a second plasmid. It is possible that either of these solutions may reduce host cell growth rates and yields or it might decrease the genetic stability of the expression system (Balbas and Bolivar 1990). Addition of inducing agents, such as indoleacrylic acid or isopropylthiogalactoside is expensive and they must be distributed quickly and evenly throughout the fermenter. Also, care must be taken to ensure that they are completely removed from the product at the processing stage, since their concentration can be relatively high, they are not metabolized and it is possible that they could co-purify with recombinant proteins.

Alternative expression systems

Alternative expression vectors that include promoters the control of which may be more tractable to the cost-effective production of recombinant proteins on an industrial scale have been developed in recent years. The promoter should fulfil the criteria listed in Table 2. Ideally, activation should occur after minor adjustments have been made to the fermentation conditions, for example by altering the flow of oxygen to the culture or by taking advantage of nutrient limitations, such as carbon, phosphorus or nitrogen source depletion, which occur during the fermentation process. The next sections describe several promoters exhibiting metabolic control that have been examined for their capacity to provide possible alternatives to established systems. The characteristics of these promoters are summarized in Table 4 and their mode of regulation is shown in Fig. 2. In some cases the efficacy of the

Table 4 Expression systems under metabolic control

Promoter	Induction principle	Induction ratio	Yield (%) of total cellular protein	References
phoA alkaline phosphatase	Phosphate deprivation	> 1000-fold	20–60	Wanner 1993; Carter et al. 1992
ugp sn-glycerol-3-phosphate transport operon	Phosphate deprivation	~ 100-fold	50	Wanner 1993; Su et al. 1990, 1991; Kasahara et al. 1991; Jarsch unpublished
araB arabinose operon	Glucose depletion and arabinose addition	1200-fold	15	Cagnon et al. 1991; Lobell and Schleif 1991
mgl methyl galactoside transport operon	Glucose addition	~ 100-fold	> 50	Schumacher et al. 1988; Müller 1989; Death and Ferenci 1994; Jarsch, unpublished
vhb Vitreoscilla haemoglobin gene	Microaerobiosis	30-fold	15	Khosla and Bailey 1988; Dikshit et al. 1990
nirB nitrite reductase operon	Anaerobiosis	100-fold	> 30	Charles et al. 1992; Chatfield et al. 1992; Schroeckh et al. 1992
pfl pyruvate formate-lyase operon	Anaerobiosis	25- to 30-fold	> 40	Sawers 1993; Oxer et al. 1991

promoter for use in an expression vector has been tested for only one or two products. However, this suffices to give an overall impression of the qualities and limitations of each and their potential for future development.

Promoters

Nutrient-regulated promoters

The gene encoding alkaline phosphatase (*phoA*) is expressed in *E. coli* at very high levels when cells are starved of inorganic phosphate (Wanner 1993). Expression can be induced more than 1000-fold with PhoA attaining levels of up to 6% of the total cellular protein from a chromosomal copy of the *phoA* gene. The *phoA* promoter is regulated both positively and negatively by the PhoB protein in response to alterations in the inorganic phosphate concentration (Table 4, Fig. 2) (Wanner et al. 1988). PhoB is the regulator component of a two-component signal-transduction cascade and, when the inorganic phosphate concentration [P_i] is above 5 mM, PhoB binds to the promoter and prevents transcription. If [P_i] drops below 1 mM PhoB becomes phosphorylated by the PhoR histidine kinase, which converts PhoB protein into a transcriptional activator. Specific details of Pho regulon control have been reviewed recently (Wanner 1993).

The *phoA* promoter has been used to construct phosphate-regulated expression vectors that have been successfully employed to produce recombinant proteins. Two examples include the production of human epidermal growth factor (Oka et al. 1985) and humanized Fab' fragments (Carter et al. 1992). In the latter study the concentration of Fab' secreted into the culture medium attained levels of 1–2 $g l^{-1}$.

The promoter from the *ugpBAECQ* operon, encoding an *sn*-glycerol-3-phosphate transport system, is also activated by PhoB phosphate in response to phosphate starvation (Wanner 1993; Su et al. 1990). Maximal promoter activation is also dependent on cAMP receptor protein (CRP) (Kasahara et al. 1991; Su et al. 1991). Reduction in both the phosphate and glucose concentration is therefore required to activate *ugp* promoter transcription. A pBR322-based expression vector has been developed (Su et al. 1990) which is applicable to large-scale fermentations (Table 4). The *ugp* promoter is very strong and exhibits both equivalent strength, under inducing conditions, and induction ratios to those of the *tac* promoter.

Phosphate-regulated promoters are essentially silent at high phosphate concentrations and are thus useful for expression of genes with products that may be toxic to *E. coli* at high concentration. Response to phosphate starvation is rapid, with maximal induction occurring within 30–60 min after the onset of nutrient limitation. The phosphate concentration at which the promoters become active (< 1 mM) is sufficient to maintain essential cellular processes permitting high-level protein synthesis (Table 4). Furthermore, their activity can be readily manipulated by limiting the P_i concentration (and glucose concentration for the *ugp* promoter) during the fermentation process. Their use should be applicable to the production of any of the categories of recombinant protein listed in Table 1.

Carbon-source-regulated promoters

The *mgl* promoter/operator region of the *mglBAEC* operon, which encodes a galactose transport system (Benner-Luger and Boos 1988), has been used to develop a series of expression vectors (Schumacher et al. 1988). Promoter activity is repressed by a protein encoded by the divergently transcribed *mglD* gene (Benner-Luger and Boos 1988) (Fig. 2). Repression is relieved by addition of D-fucose. The promoter is also

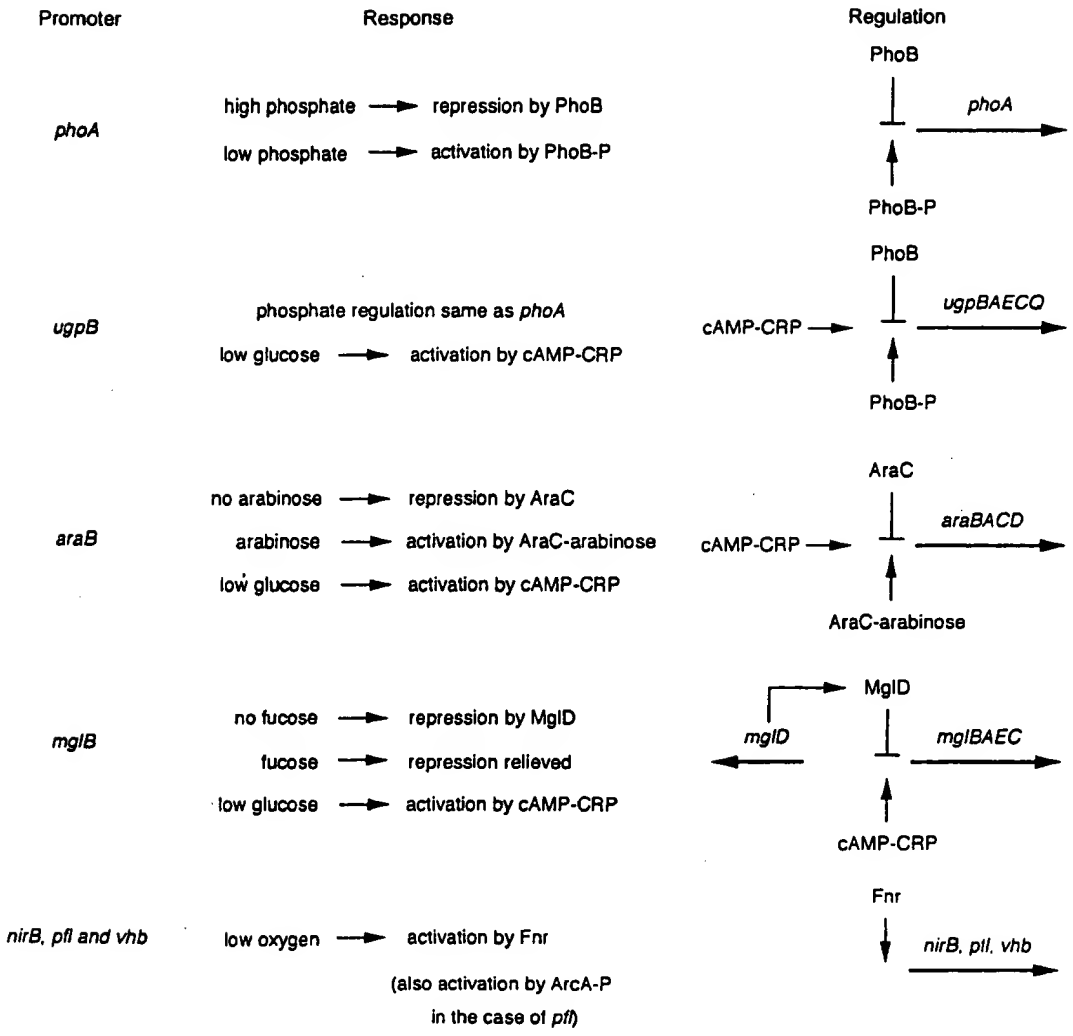


Fig. 2 Summary of the regulation of metabolically controlled promoters. → Genes or operons; ↑ activation; ⊥ repression of promoter activity

dependent on the cAMP-CRP complex for maximal activation and is therefore subject to glucose repression. Construction of a host strain and vector with the *mglD* gene deleted obviates the requirement for fucose addition and delivers an expression system that is controlled solely by catabolite repression. Induction of promoter activity is approximately 100-fold and this promoter has been used in an expression system that produced recombinant antigen from *Echinococcus multilocularis* at levels of 5 $g l^{-1}$ (Müller et al. 1989) and more recently the glucose dehydrogenase from *Acinetobacter calcoaceticus*, also at concetrations of several grams per litre (M. Jarsch, unpublished results). Obviously the expression system regulated solely by glucose repression has the advantage that it can be readily and cheaply controlled in a large-scale process but it has the disadvantage that it can only be used for production of recombinant proteins that are both nontoxic to *E. coli* and have a background synthesis that does not reduce growth rates or cell yields extensively. Significantly, a recent study (Death and Ferenci 1994) has reported that when *E. coli* cells are grown in conditions of controlled glucose limitation sufficient endogenous inducer is synthesized to induce *mgl* operon expression to high levels. Optimization of the growth conditions should therefore permit the inclusion of the *mglD* gene in the vector, yielding an expression system with greater flexibility and concomitantly minimizing production costs by obviating the requirement for exogenous inducer.

A second sugar-regulated expression system utilizes the *araB* promoter from the arabinose operon of *Salmonella typhimurium* (Cagnon et al. 1991). The

activity of the *araB* promoter is subject to dual regulation through catabolite repression and induction by arabinose (Table 4). Promoter activity is controlled by the Ara C and CRP transcription factors (Fig. 2). Ara C represses transcription in the absence of arabinose but is a positive regulator of the promoter when it is itself complexed with D-arabinose (Lobell and Schleif 1991). D-Fucose is a structural analogue of arabinose and competes very efficiently with it to increase repression. Promoter activation, however, depends completely on the cAMP-CRP complex, even when arabinose is present (Lobell and Schleif 1991). It is possible to regulate the promoter over a 1200-fold range and vectors have been constructed that give yields of recombinant protein in the range of 30% of the total cellular protein (Cagnon et al. 1991). This system is similar to *lac*-promoter-based expression vectors in that glucose depletion and supply of an inducer form the basis of the regulation principle. The *araB* promoter has the added advantage that it can be switched off almost completely by fucose supplementation and hence provides a slightly more flexible system, in particular for production of toxic proteins. Moreover, arabinose is more cost-effective as an inducer compared with IPTG; on the basis of current market prices, IPTG is twice as expensive as D-arabinose if it is assumed that they are used at a final concentration of 5 mM and 1% (w/v) respectively.

Oxygen-regulated promoters

Promoters having an activity that can be modulated by varying the dissolved oxygen concentration of the culture medium offer several favourable advantages for the design of vectors to be used in the industrial production of recombinant proteins. First, expression is completely independent of the growth medium. Second, no addition of any inducing agents is necessary. Third, there is no dependence on a particular *E. coli* host strain. Fourth, the promoters are functional in the late exponential or early stationary phase of growth. Finally, high dissolved O_2 concentrations can be maintained until high cell densities are attained, and simply controlling the rate of aeration, the proportion of the oxygen in the gas mixture or regulating the speed of stirring can induce the expression of the recombinant gene. Since ensuring adequate aeration in high-density cultures is in any case a problem in fermenters, the use of oxygen-regulated promoters provides an inexpensive means of controlling product synthesis.

Three expression systems have been developed independently but all are based on the same regulation principle (Table 4, Fig. 2). The first system uses the promoter from the gene encoding bacterial haemoglobin (VHb). VHb is produced in large amounts by the gram-negative obligate aerobe *Vitreoscilla* when the dissolved O_2 becomes limiting (Khosla and Bailey 1989). The protein appears to function in oxygen delivery to the vigorously respiring membranes of the organism and it has been demonstrated that expression of the *vhb* gene in *E. coli* increased both the rate of growth and the cell densities attained (Khosla and Bailey 1988). The activity of the *vhb* promoter is regulated by oxygen in *E. coli* and it is likely that this may be controlled by the oxygen-responsive transcription factor Fnr (Spiro and Guest 1990). In a two-stage batch fermentation Khosla et al. (1990) overproduced chloramphenicol acetyltransferase and β-galactosidase, in independent constructs, to levels approaching 10% of the soluble cellular protein by growing recombinant *E. coli* to high cell densities at a dissolved O_2 concentration of 20% and then reducing the O_2 concentration to below 5% air saturation, which induced expression. An overall 30-fold increase in promoter activity was achieved (Khosla et al. 1990). Because the activity of the *vhb* promoter is optimal under conditions of microaerobiosis this means that the dissolved O_2 concentration must be carefully controlled to elicit high-level expression; the promoter has reduced activity when the culture becomes anaerobic (Khosla and Bailey 1989). It is noteworthy, however, that this promoter is functional over a broad host range, which includes *Pseudomonas*, *Azotobacter* and *Rhizobium* species (Dikshit et al. 1990).

We have developed an expression vector using the pyruvate formate-lyase operon regulatory region (Sawers and Böck 1989; Böck et al. 1990). Pyruvate formate-lyase is the central enzyme of anaerobic catabolism in *E. coli* and can reach levels exceeding 3% of the total cellular protein. The *pfl* gene is expressed anaerobically from multiple promoters and the induction ratio is maximally 30-fold (Table 4) (Sawers and Böck 1989). Anaerobic regulation is mediated by the Fnr and ArcA transcription factors (Sawers 1993). The vector that has been constructed incorporates two of the promoters plus the complete regulatory region and has been used successfully to produce in *E. coli* high levels of cholesterol oxidase from *Brevibacterium sterolicum* (Table 4). A disadvantage of this expression system is that the promoter has a significant basal activity when the organism is grown at high dissolved O_2 concentrations, therefore it is unsuitable if low background promoter activity is desired or necessary. Although the induction ratios of both the *pfl* and the *vhb* systems are moderate compared with the others listed, nevertheless they are based on strong promoters that allow high-level heterologous gene expression and their ease of control makes them very cost-effective.

Finally, a portion of the Fnr-dependent *nirB* promoter from the anaerobically inducible nitrite reeductase operon of *E. coli* has been used to produce an effective expression vector (Oxer et al. 1991). Like the *pfl* promoter *nirB* promoter activity is dependent on the Fnr transcription factor but it is also induced by nitrite. Oxer and colleagues (Oxer et al. 1991) removed the nitrite-responsive regulatory sequences but left the

recognition sequences for the Fnr protein intact. The resulting derivative exhibited tight transcriptional control in fermenter studies (Table 4) (Charles et al. 1992; Chatfield 1992). Two examples of heterologously produced proteins under the control of the *nirB* promoter have been described: tetanus toxin fragment C was produced at levels of 20% total cellular protein while the *Bordetella pertussis* antigen pertactin was synthesized at levels exceeding 30% of the total cellular protein (Oxer et al. 1991). A very low level of expression was observed when the oxygen supply was maintained, whereas anaerobiosis caused an approximately 100-fold induction of activity.

Conclusions and perspectives

The support of fundamental research has started to pay dividends in providing industry with new, strong, regulated expression systems that are suited to the production of recombinant proteins on a large scale in fermentation processes. Although they may not replace the established expression vectors for all purposes they offer attractive alternatives that have comparable promoter strength and, in the majority of cases, are easier and cheaper to control. The alternative promoters described, despite being strong, are not ferocious and experience using these systems to date has shown that not only is the yield of recombinant product high but also the fraction of insoluble, functionally inactive protein is generally quite low.

It is likely that future research will reveal further metabolically regulated promoters that may be suitable for the development of expression vectors. Indeed, a patent has been registered (Schroeckh et al. 1992) in which an expression vector has been developed that includes the *glnAp2* promoter from the glutamine synthetase operon of *E. coli*. The activity of the promoter is controlled by RNA polymerase containing an alternative sigma factor and is positively regulated by nitrogen deprivation. Alternative sigma factors confer upon RNA polymerase holoenzyme the ability to transcribe specifically promoters that have a different recognition sequence compared with the general sigma factor. A further class of regulated promoters, recognized by an alternative sigma factor, that could potentially be exploited for designing expression vectors includes those the activity of which is switched on only in the stationary phase of growth (Hengge-Aronis 1993). Finally, a series of vectors has recently been developed based on the *proU* promoter that is regulated in response to osmolarity (Herbst et al. 1994). The activity of the promoter is controlled by varying the concentration of NaCl in the growth medium and it exhibits excellent regulatory properties. These vectors therefore should prove extremely useful for the cost-effective production of recombinant proteins.

These promoters provide the basis for development of further novel expression vectors and, as our knowledge of the regulatory mechanisms underlying their control advances, it will perhaps be possible to combine features from different systems to develop hybrid or synthetic promoters with improved regulatory characteristics. The feasibility of this is shown by the development of hybrid promoters combining properties of established systems to deliver very efficient, tightly regulated expression vectors (Wülfing and Plückthun 1993; Chen et al. 1993).

Acknowledgements The work carried out by G.S. was performed in the Department of Microbiology at the University of Munich and was supported by the Deutsche Forschungsgemeinschaft.

References

Balbas P, Bolivar F (1990) Design and construction of expression plasmid vectors in *Escherichia coli*. Methods Enzymol 185:14–37

Beckwith J (1987) The lactose operon. In: Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umbarger HE (eds) *Escherichia coli* and *Salmonella typhimurium*: molecular and cellular biology. American Society for Microbiology, pp 1444–1452

Benner-Luge D, Boos W (1988) The *mglB* sequence of *Salmonella typhimurium* LT2; promoter analysis by gene fusions and evidence for a divergently oriented gene coding for the *mgl* repressor. Mol Gen Genet 214:579–587

Böck A, Sawers G, Jarsch M, Herbst R (1990) Rekombinante DNA und Expressionsvektor. European patent application 0393684

Brosius J, Erfle M, Storella J (1985) Spacing of the −10 and −35 regions in the *tac* promoter. J Biol Chem 260:3539–3541

Cagnon C, Valverde V, Masson JL (1991) A new family of sugar-inducible expression vectors for *E. coli*. Protein Eng 4:843–847

Carter P, Kelley RF, Rodrigues ML, Snedecor B, Covarrubias M, Velligan MD, Wong WLT, Rowland AM, Kotts CE, Carver ME, Yang M, Bourell JH, Shepard, HM, Henner D (1992) High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. Biotechnology 10:163–167

Charles IG, Makoff AJ, Oxer MD, Peakman TC, Cole JA (1992) Production of recombinant proteins. Patent WO 9215688

Chatfield SN, Charles IG, Makoff AJ, Oxer MD, Dougan G, Pickard D, Slater D, Fairweather NF (1992) Use of the *nirB* promoter to direct the stable expression of heterologous antigens in *Salmonella* oral vaccine strains – development of a single-dose oral tetanus vaccine. Biotechnology 12:888–892

Chen W, Kallio PT, Bailey JE (1993) Construction and characterization of a novel crossregulation system for regulating cloned gene expression in *Escherichia coli*. Gene 130:15–22

Death A, Ferenci T (1994) Between feast and famine: endogenous inducer synthesis in the adaptation of *Escherichia coli* to growth on limiting carbohydrates. J Bacteriol 176:5101–5107

DeBoer HA, Comstock LJ, Vasser M (1983) The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc Natl Acad Sci USA 80:21–25

Dikshit KL, Dikshit RP, Webster DA (1990) Study of *Vitreoscilla* globin (*vhb*) gene expression and promoter activity in *E. coli* through transcriptional fusion. Nucleic Acids Res 18:4149–4155

Fuller F (1982) A family of cloning vectors containing the *lacUV5* promoter. Gene 19:43–54

Gold L, Stormo GD (1990) High-level translation initiation. Methods Enzymol 185:89–93

Gottesman S (1990) Minimizing proteolysis in *Escherichia coli*: genetic solutions. Methods Enzymol 185:119–129

Gralla JD (1990) Promoter recognition and mRNA initiation by *Escherichia coli* $E\sigma^{70}$. Methods Enzymol 185:37–54

Gross G (1989) High-level gene expression in *Escherichia coli*. Chimicaoggi March:21–29

Hengge-Aronis R (1993) Survival of hunger and stress: the role of *rpoS* early stationary phase gene regulation in *E. coli*. Cell 72:165–168

Herbst B, Kneip S, Bremer E (1994) pOSEX: vectors for osmotically controlled and finely tuned gene expression *Escherichia coli*. Gene 151:137–142

Hockney RC (1994) Recent developments in heterologous protein production in *Escherichia coli*. Trends Biotechnol 12:456–463

Kasahara M, Makino K, Amemura M, Nakata A, Shinagawa H (1991) Dual regulation of the *ugp* operon by phosphate and carbon starvation at two interspaced promoters. J Bacteriol 173:549–558

Khosla C, Bailey JE (1988) Heterologous expression of a bacterial hemoglobin improves the growth properties of recombinant *Escherichia coli*. Nature 331:633–635

Khosla C, Bailey JE (1989) Characterization of the oxygen-dependent promoter of the *Vitreoscilla* hemoglobin gene in *Escherichia coli*. J Bacteriol 171:5995–6004

Khosla C, Curtis JE, Bydalek P, Swartz JR, Bailey JE (1990) Expression of recombinant proteins in *Escherichia coli* using an oxygen-responsive promoter. Biotechnology 8:554–558

Kopetzki E, Lehnert K, Buckel P (1994) Enzymes in diagnostics: achievements and possibilities of recombinant DNA technology. Clinical Chem 40:688–704

Lobell RB, Schleif RF (1991) AraC-DNA looping: orientation and distance-dependent loop breaking by the cAMP receptor protein. J Mol Biol 218:45–54

Makoff AJ, Oxer MD (1991) High level expression in *E. coli* using mutant forms of the *lac* promoter. Nucleic Acids Res 19: 2417–2421

Missiakis D, Georgopoulos C, Raina S (1993) Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds in vivo. Proc Natl Acad Sci USA 90:7084–7088

Müller N, Vogel M, Gottstein NB, Scholle A, Seebeck T (1989) Plasmid vector for overproduction and export of recombinant protein in *Escherichia coli*: efficient one-step purification of a recombinant antigen from *Echinococcus multilocularis* (Cestoda). Gene 75:329–334

Nygren P-A, Ståhl S, Uhlen M (1994) Engineering proteins to facilitate bioprocessing. Trends Biotechnol 12:184–188

Oka T, Sakamoto S, Miyoshi K, Fuwa T, Yoda K, Yamasaki M, Tamura G, Miyake T (1985) Synthesis and secretion of human epidermal growth factor by *Escherichia coli*. Proc Natl Acad Sci USA 82:7212–7216

Oxer MD, Bentley CM, Doyle JG, Peakman TC, Charles IG, Makoff IJ (1991) High level heterologous expression in *E. coli* using the anaerobically-activated *nirB* promoter. Nucleic Acids Res 19:2889–2892

Remaut E, Stanssans P, Fiers W (1981) Plasmid vectors for high-efficiency expression controlled by the p_L promoter of coliphage lambda. Gene 15:81–93

Richter G (1983) Glucose oxidase. In: Godfrey T, Reichelt J (eds) Industrial enzymology. The application of enzymes in industry. MacMillan, New York, pp 428–436

Rosenberg M, Ho YS, Shatzman A (1983) The use of pKC30 and its derivatives for controlled expression of genes. Methods Enzymol 101:123–138

Santos MAS, Tuite MF (1993) New insights into mRNA decoding–implications for heterologous protein synthesis. Trends Biotechnol 11:500–505

Sawers G (1993) Specific transcriptional requirements for positive regulation of the anaerobically inducible *pfl* operon by ArcA and FNR. Mol Microbiol 10:737–747

Sawers G, Böck A (1989) Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. J Bacteriol 171:2485–2498

Schroeckh V, MacCarthy JEG, Deckwer WD, Riesenberg D (1992) Expressionsvektor sowie Verfahren zur Expression rekombinater Gene unter Stickstoffmangelbedingungen. German patent DE 4102594

Schumacher G, Jarsch M, Boos W (1988) Expressionsvektor zur regulierbaren Expression von Fremdgenen in Prokaryonten. German patent application DE 3716957

Spiro S, Guest JR (1990) FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol Rev 75:399–428

Squires CL, Lee FD, Yanofsky C (1975) Interaction of *trp* repressor and RNA polymerase with the *trp* operon. J Mol Biol 92:93–111

Stark MJR (1987) Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. Gene 51:255–267

Studier FW, Moffat BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. J Mol Biol 189:113–130

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol 185:60–89

Su T-Z, Schweizer H, Oxender DL (1990) A novel phosphate-regulated expression vector in *Escherichia coli*. Gene 90:129–133

Su T-Z, Schweizer H, Oxender DL (1991) Carbon-starvation induction of the *ugp* operon, encoding the binding protein-dependent *sn*-glycerol-3-phosphate transport system in *Escherichia coli*. Mol Gen Genet 230:28–32

Tabor S, Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci USA 82:1074–1078

Tacon W, Carey N, Emtage S (1980) The construction and characterization of plasmid vectors suitable for the expression of all DNA phases under the control of the *E. coli* tryptophan promoter. Mol Gen Genet 177:427–438

Wanner BL (1993) Gene regulation by phosphate in enteric bacteria. J Cell Biochem 51:47–54

Wanner BL, Wilmes MR, Young DC (1988) Control of bacterial alkaline phosphatase synthesis and variation in an *Escherichia coli* K-12 *phoR* mutant by adenyl cyclase, the cyclic AMP protein, and the *phoM* operon. J Bacteriol 170:1092–1102

Wülfing C, Plückthun A (1993) A versatile and highly repressible *Escherichia coli* expression system based on invertible promoters: expression of a gene encoding a toxic product. Gene 136: 199–203

Yansura DG, Henner DJ (1990) Use of the *Escherichia coli trp* promoter for direct expression of proteins. Methods Enzymol 185:54–60

GENE TRANSFER TO CEREALS: AN ASSESSMENT

Ingo Potrykus
Institute for Plant Sciences, Swiss Federal Institute of Technology (ETH), ETH Zentrum, CH-8092 Zürich, Switzerland.

In the following review, I present an assessment of the realities and possibilities of effecting gene transfer to cereal crops. I discuss why *Agrobacterium* has been unsuccessful with cereals, what alternatives have been tested, the extent to which they have yielded transgenic plants, and their potential agronomic utility. The discussion, necessarily subjective, is framed within a rigid definition of what constitutes proof of gene integration, and the biological factors affecting transformation competence.

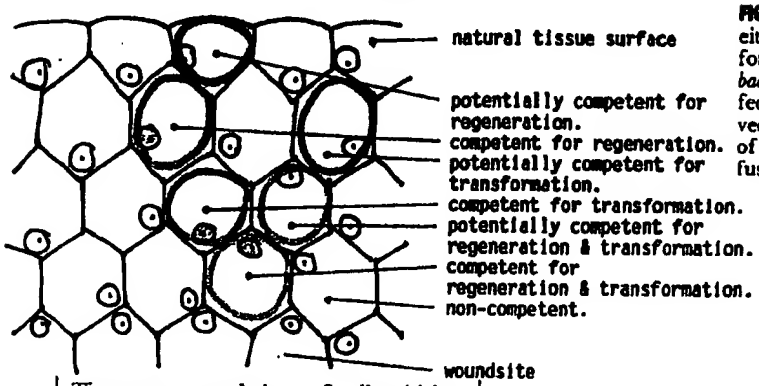
The first transgenic plants expressing engineered foreign genes where recovered in 1984 (see review[1]). The five years since then have yielded such exciting results it is not surprising that numerous optimistic reviews have been written. One of the most recent[1] makes the following typical statement: "Dramatic progress has been made in the development of gene transfer systems for higher plants. . . . In view of the rapid progress that is being made, it is likely that all major dicotyledoneous and monocotyledoneous crop species will be amenable to crop improvement by genetic engineering within the next few years." Such optimism is understandable and generally good for further progress, but my personal experience in working towards the genetic engineering of cereals for the last 18 years convinces me that we still have serious problems in front of us. Success requires more than occasional gene transfer into experimentally well suited varieties of some species. It requires routine and efficient gene transfer into any desired variety of any given species, changing the genome only by the addition of one defined gene. It seems to me that we are really not yet close to such a situation. Usually gene transfer to plants is achieved with *Agrobacterium*, but as this biological vector does not function with cereals a variety of alternative approaches for gene transfer to these species have been developed, of which only one has so far been successful. Even the successful method has its problems. In the following review I will discuss why *Agrobacterium* has been unsuccessful with cereals, what alternative methods have been tested, which of those have yielded transgenic cereals, which may have a realistic chance to become a successful technique, and which may not have much potential at all. The assessment will be subjective. It will be based on a rigid definition of what constitutes *proof* of successful integrative transformation. Those who disagree with the view that indicative evidence is misleading may not agree with this assessment. The review will also be based on an interpretation of the biological factors affecting gene transfer, and several statements will be made for which no solid experimental data are available. Acceptance or refusal of these statements will not affect assessment of the available data but will only influence attitudes as to the future potential of the various approaches.

1. Proof and competence (Figure 1). If one trusts indicative evidence, there are numerous different methods to produce transgenic cereals. However, since no transgenic cereals exist except for those that have been recovered from protoplasts and direct gene transfer, there must be something wrong with the indicative evidence. Indeed many researchers have obviously been misled by artifacts and it is good advice to believe in transgenic plants only if suitable proof is available. What constitutes suitable proof? Neither genetic, phenotypic or physical data alone are acceptable. *Proof* for integrative transformation requires:

(1) Serious controls for treatments and analysis. (2) A

8(6):535-542

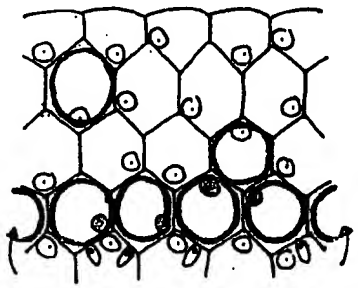
1 TISSUES ARE POPULATIONS OF CELLS NON COMPETENT, POTENTIALLY COMPETENT, AND COMPETENT FOR REGENERATION & TRANSFORMATION.



Tissues are populations of cells which are either non-competent, potentially competent, or competent for regeneration and/or integrative transformation.

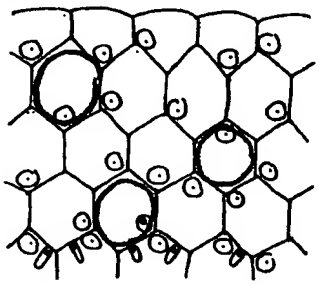
FIGURES 1–11 (1) Tissues are populations of cells which are either non-competent, potentially competent, or competent for regeneration and/or integrative transformation. (2) *Agrobacterium* and dicots. (3) *Agrobacterium* and cereals. (4) Agroinfection and dicots. (5) Agroinfection and cereals. (6) Viral vectors. (7) Incubation of dry seeds/embryos. (8) Incubation of tissues/cells in DNA. (9) Pollentube pathway. (10) Liposome fusion. (11) Liposome injection.

2



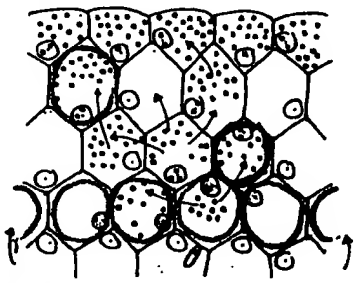
Wound response makes potentially competent cells competent

3



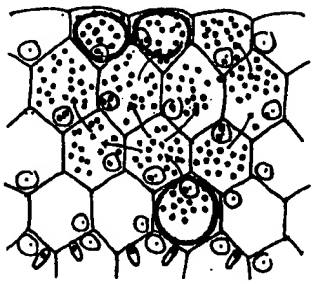
No competent cells because wound response is missing.

4



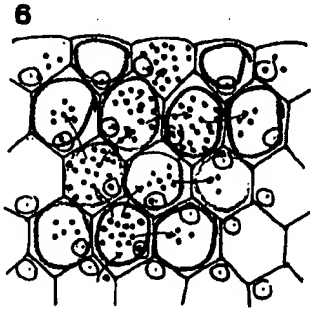
Virus spreads systemically, but does not integrate.

5



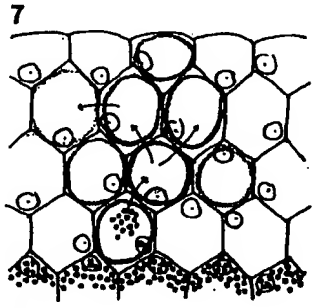
Virus spreads systemically, but does not integrate.

6



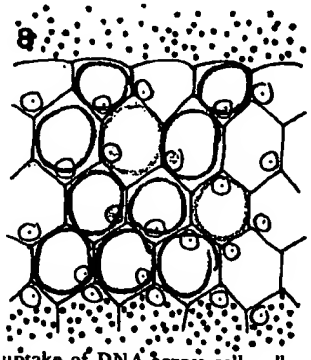
Systemic spread, no integration, no transmission to offspring.

7



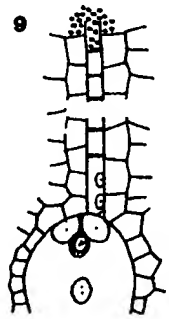
Viral DNA spreads, does not integrate. Non-viral DNA does not enter tissue.

8



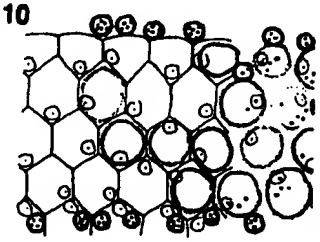
No uptake of DNA across cell walls, or as very rare event only

9



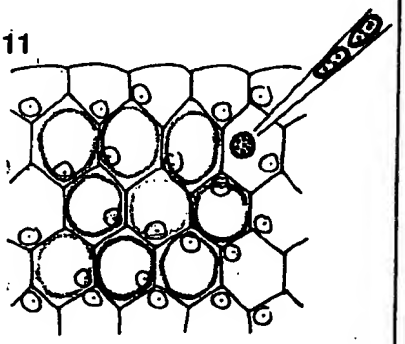
No open pipes to the egg cell; DNA adhesion to cell walls; nucleases.

10



Transfers DNA into protoplasts but not into cells with cell walls.

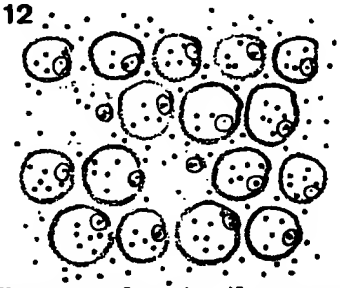
11



Transfers DNA into cells avoiding the vacuole.

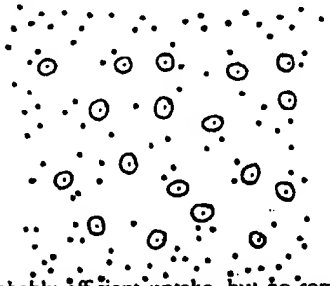
FIGURES 12–23 (12) Protoplasts and direct gene transfer. (13) Protoplasts from cereal plants. (14) Microlaser. (15) Electrophoresis into tissues. (16) Biolistics or particle gun. (17) Biolistics and cell cultures. (18) Microinjection into tissues. (19) Microinjection into proembryos. (20) *Agrobacterium* and zygotic proembryos. (21) Macroinjection. (22) Pollen incubation. (23) Electroporation.

12



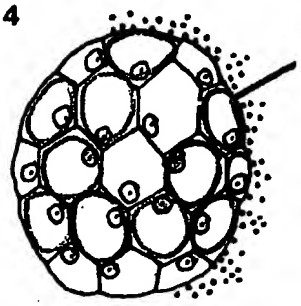
Efficient transformation if competent protoplasts available.

13



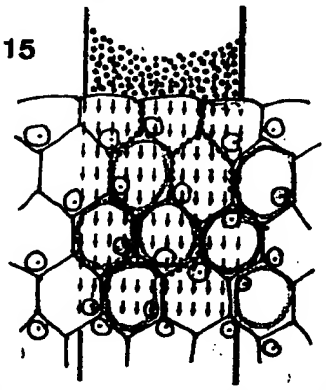
Probably efficient uptake, but no competent protoplasts.

14



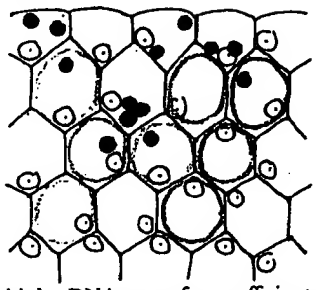
Opens holes in cell walls and membranes; DNA adhesion to cell walls.

15



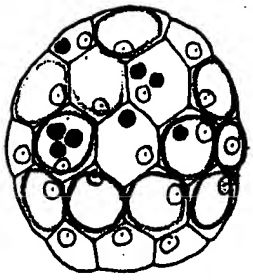
Probably no DNA transport across walls, no transgenic clones so far.

16



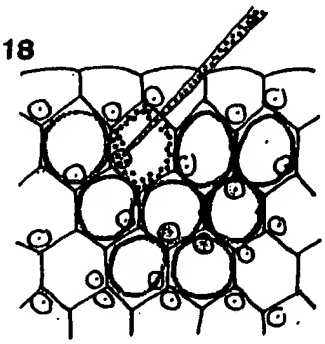
Multiple DNA transfers; efficient in transient, inefficient in integrative transformation.

17



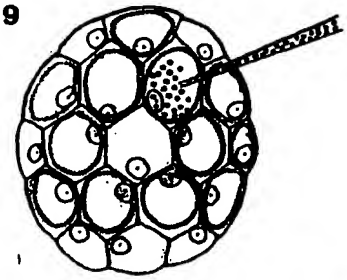
Transgenic clones at low frequency; no transgenic cereals, but promising.

18



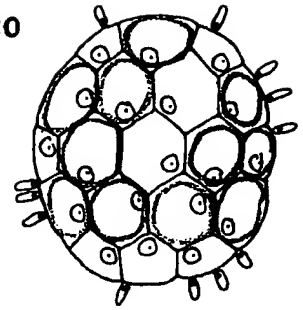
Precise delivery of DNA quantities; Transgenic clones from protoplasts.

19



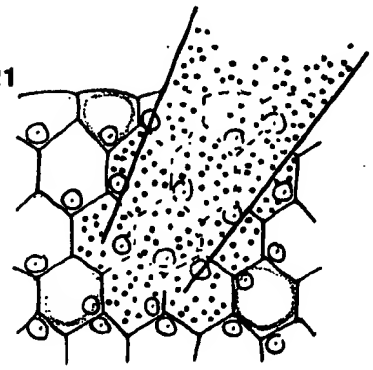
Transgenic chimeras at low frequency; no transgenic cereals, but promising.

20



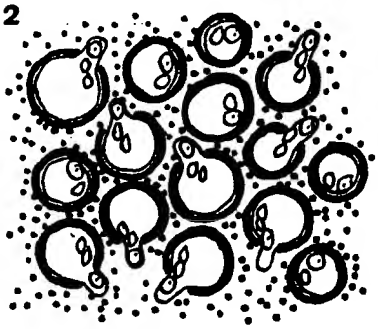
So far no transgenic tissue even in tobacco.

21



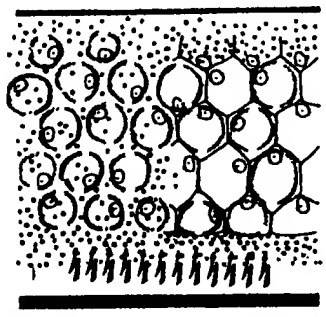
Injection destroys cells which receive DNA; no DNA transport to competent cells.

22



Probably no uptake across pollentube wall; nucleases; no integration into sperm cells.

23



Efficient DNA transfer to protoplasts. No transport across cell walls.

tight correlation between treatment and predicted results. (3) A tight correlation between physical (Southern blot, *in situ* hybridization) and phenotypic (enzyme assays) data. (4) *Complete Southern analysis* containing (a) the predicted signals in high molecular weight DNA, in hybrid fragments between host DNA and foreign gene, and the complete gene, and (b) evidence for the absence of contaminating DNA fragments or the identification of such fragments. (5) Data that allow discrimination between false positives and correct transformants in the evaluation of the phenotypic evidence. (6) Correlation of the physical and phenotypic evidence with transmission to sexual offspring, as well as genetic and molecular analysis of offspring populations.

Consideration of the *biology* of gene transfer may be helpful in understanding why some techniques work and others never do. It may also help in assessing the future potential of the various approaches. A transgenic plant can only result from integrative transformation in a totipotent cell or a cell that has a clonal connection to the "germline".

(1) Not all plant cells are totipotent. (2) Plant cells differ in their capacity to respond to triggers, a phenomenon termed *competence.*) (3) Cells from which it is hoped to regenerate transgenic plants must be competent for both regeneration (in a broad sense) and integrative transformation. (4) Plant tissues are composed of cells competent for many different responses. Considering the two states of competence essential for recovery of transgenic plants the following situation has to be considered: (a) A very small minority of cells in plant tissues will be *competent for both transformation and regeneration.* (b) Others will be competent for transformation *or* regeneration. (c) A larger fraction of the cell population will be *potentially competent,* which means that given the correct treatment they will have the potential to shift to the competent state. (d) A variable proportion of cells will not even be potentially competent but will be *non-competent.* (5) The relative composition of cell populations in tissues is determined by the genotype, the type of organ, the developmental state of the organ, and even the individual history of the experimental plant. (6) The most effective trigger for shifting potentially competent cells to the competent state is mechanical (and enzymatic) wounding. The wound response[2] is probably the biological basis of regeneration from somatic cells. (7) Plant species differ in their wound response as do different tissues of the same plant. Graminaceous plant species, especially the cereals and maize, have only a very rudimentary or no wound response. (8) For some genotypes it is possible to proliferate cells competent for regeneration under conditions that maintain this state[3]. Such cell cultures contain cells competent for regeneration and (after protoplasting?) competent for integrative transformation. (9) Plant cell walls are efficient barriers and traps for DNA molecules. (10) Genes can be transported into cells across cell walls with the help of *Agrobacterium*, "biolistics" and microinjection. (11) Production of transgenic plants requires efficient gene transfer into cells competent for regeneration and integrative transformation. (12) Competence for integrative transformation is obviously very different from competence for transient expression. (13) Non-viral DNA can integrate into the host genome. Its presence in a cell does not guarantee its integration. (14) Non-viral DNA does not move from cell to cell but is restricted to the cell to which it has been delivered. (15) Viral DNA (and RNA) moves from cell to cell and can spread systemically throughout an entire plant. It is, however, probably excluded from the meristems and the "germline". (16) Viral DNA does not integrate into the host genome even if present at very high copy number.

On this basis an assessment of the various gene transfer approaches presented in the literature or discussed at international meetings is relatively easy. The cartoons help understanding by visualising the biological problems. As the background literature is extensive, I have relied to a large extent on recent reviews. I intend this to be a provocative and working hypothesis. It is, however, in agreement with all the available data from the literature and from public meetings.

2. *Agrobacterium* and dicots[4-7] (Fig. 2). *A routine and efficient method for the production of transgenic plants from numerous non-cereal species.* This includes many cases where transformation was possible only on the basis of intensive screening for the optimal combination of plant genotype and bacterial strain, or where transformation was possible only at a short developmental stage of specific tissues of the host plant, and in addition, numerous cases where transformation has so far not been possible. Interpretation: Plants and tissues differ in their wound response. Only plants and tissues with a pronounced wound response develop larger populations of wound-adjacent competent cells for efficient transformation. Dicots that have not been transformed probably do not show the appropriate wound response.

3. *Agrobacterium* and cereals[8] (Fig. 3). *No transgenic cereal recovered so far; very little potential.* Most attempts have not been published because they were negative; some promising data have been presented at international meetings but since no proof was reported, the data must be considered artifactual. Transformation of "monocots"[9] (and citations therein) are of no importance in this context as cereals are not difficult to transform because they are monocots, but because they show no wound response. Monocots with a wound response (e.g. asparagus and yam) are as easy to transform as dicots with a wound response. And dicots without a wound response are as difficult to transform as cereals. Why has it been impossible to transform cereals with *Agrobacterium*? Wounding of differentiated cereal tissue does not lead to the wound response-induced dedifferentiation in wound-adjacent cells. Therefore no competent cells are available. Instead wounding leads to the death of the wound-adjacent cells. Even though *Agrobacterium* is very efficiently at transferring its T-DNA into cereal cells, (see section 5. *Agroinfection and cereals*) integration of this T-DNA can not lead to transgenic clones because the receptor cells die. It is not as easy to understand why even in those experiments where meristematic tissue (e.g. leaf base or split shoot tip), which can be induced to form proliferating cultures and plants, did not yield transgenic plants. One possibility is that cereal cell cultures are not the consequence of proliferating wound meristems but are based on adventitious meristems[10]. Wounding plus *in vitro* culture does not lead to many competent cells but to a few meristem initials that proliferate as meristems. Meristematic cells may not be transformation competent (see section 17. *Microinjection* and section 18. *Agrobacterium and zygotic proembryos*).

Raineri et al.[34] presented three different lines of evidence giving a reasonable inference of *Agrobacterium*-mediated transformation of rice (*Oryza Sativa*). Although definitive proof, as defined within the present context, is still lacking, it should be relatively straightforward to obtain the necessary data.

4. Agroinfection and dicots[11,12] (Fig. 4). *Agroinfection can lead to transgenic plants via T-DNA integration.* Viral DNA integrated into the T-DNA of the Ti-plasmid of *Agrobacterium* can be delivered into plant cells with the normal T-DNA transfer process. The consequences of agroinfection in dicots with a wound response are the

following: The virus enters the cell as part of the T-DNA. It is released to form a functional virus that replicates and spreads systematically. It may not be necessary for the T-DNA to integrate in order to release the virus. Systemic spread of the virus as a consequence of agroinfection is, therefore, no proof of integration of foreign DNA. T-DNA *can*, however, integrate and thus agroinfection can lead to integration of viral DNA in the wound-adjacent cells and consequently to transgenic plants containing integrated viral DNA. This is not different from normal T-DNA transformation. The important point is that there is no integration of the systemically spreading virus. Can agroinfection be used to transform cereals?

5. Agroinfection and cereals[13] (Fig. 5). *This method has very little potential for the production of transgenic cereals.* Agroinfection has been shown to lead to systemic spread of maize streak virus. This showed for the first time that *Agrobacterium* can transfer its T-DNA to cereal cells. Later it was demonstrated that the efficiency of transfer is comparable to dicot systems. Does this method, therefore, have potential for the production of transgenic cereals? In cereals, agroinfection leads to the transfer of the virus-carrying T-DNA into wound-adjacent cells. The virus is released, replicates and spreads systemically. If it reaches rare competent cells somewhere in the plant body it will not integrate. At the woundsite, the T-DNA faces the same problems as discussed in section 3. Thus the chances that agroinfection will produce transgenic cereals are minimal and no different normal *Agrobacterium* infection—unless somebody finds a way to induce integration of viral DNA, or of foreign DNA integrated into a replicating and spreading virus, as could be envisaged for a transposable element carried through the plant by a spreading virus.

6. Viral vectors[14–17] (Fig. 6). *This method has very little potential for production of transgenic cereals, although it is very interesting for amplification of genes and gene products.* Viruses spread systemically throughout the plant from the infection site and can replicate many thousands of copies per cell. The discovery that RNA virus genomes can be reverse transcribed to yield cDNA clones that again are infective opened the possibility of using genetic engineering technology not only with the relatively small group of DNA viruses but also with the larger group of RNA viruses. According to the available evidence, viruses do not integrate into the host genome and they are excluded from meristems and thus from transmission to sexual offspring. These facts determine the advantages and disadvantages of viruses as vectors for possible genetic engineering in plants. Although it would be very difficult (except e.g. via integrated transposable elements) to use viral vectors for integrative transformation, they have invaluable potential for gene amplification and systemic spread within individual plants.

7. Incubation in DNA of dry seeds or embryos[18–21] (Fig. 7). *Thus far, no transgenic plants have been recovered, not much potential.* Incubation of seeds in DNA has yielded indicative evidence since the 1960's. However, no proof for integrative transformation has ever been presented. Töpfer et al.[19] describe experiments in which every precaution was taken to avoid the experimental pitfalls of earlier experiments and these yielded very interesting evidence for transformation following incubation of dry seeds or embryos in engineered viral and non-viral DNA. Although the experiments contain convincing controls and clearly demonstrate the presence and expression of defined marker genes as well as the replication of engineered viral DNA, they do not provide proof for integrative transformation. Their conclusion that the data demonstrate uptake of the foreign DNA into the cells of the embryos is really one hypothesis. It is, to date not possible to exclude the alternative hypothesis that the DNA data are the result of transient reactions in the micro-environment of open cells at the large wound site and that the virus DNA data are, in addition, due to systemic spread. Only the analysis of offspring from regenerated plants will finally show which hypothesis is correct. So far, I favour the following interpretation: If engineered viral DNA is used, it may spread systemically and even reach competent cells; it will not integrate. If non-viral DNA is used, it will most probably not even be taken up into the wound-adjacent cell layer or only at very low efficiency. Wound-adjacent cells will not include competent cells or they will be present only at very low frequency, and these would not be induced to proliferate to form cell clones. The DNA would not reach the meristematic cells of the shoot apex because it would have to travel across numerous cell walls. If the treatment were to be modified to allow direct access to the shoot meristem, regeneration would not occur from the wounded shoot apex but from non-wounded axillary meristems of the leaves. Therefore regeneration of transgenic plants is highly improbable.

8. Incubation in DNA of tissues or cells (Fig. 8). *No transgenic tissues or plants have been recovered; very low potential.* There have been many approaches where seedlings, organs, tissues, cells, or cell cultures of numerous plant species have been brought into direct contact with foreign DNA and defined marker genes. Treatments also included experimental designs making use of open plasmodesmata or loosening of cell wall structures. There were also treatments ensuring that competent cells were available at a sufficient frequency. Even in experiments that would have recovered extremely rare events of integrative transformation, there is not a single proven case of integrative transformation. Experiments relying on the passage of functional genes across cell walls have very little chance of success, not only because the cell wall is a perfect barrier to large DNA molecules, but because it is also an efficient trap. Even if there were occasional transfer there are other negative parameters that act in a multiplicative way: (a) attachment to cell walls, (b) transport across further cell walls, (c) no mechanism for DNA transport, (d) competent cells have to be reached. The combination of several low frequency events will cause problems even if one step may occasionally work.

9. Pollentube pathway[22] (Fig. 9). *No transgenic plants have been recovered; probably not much potential.* If it were possible to deliver DNA to the zygote via open pollen tubes in the course of normal pollination, this would be very attractive. Unfortunately, the recent publication presenting molecular data on transgenic rice plants does not present proof. The Southern data do not show integration into high molecular weight DNA and defined hybrid fragments, and can be easily interpreted as artifacts: the dot blot technique is prone to artifacts and the enzyme data are not reliable because cereals have a rich record of false positives with the assay used. It is also difficult to envisage how this system *should* work: the pollen tubes are not open pipes but sealed off with callose plugs: the DNA will be trapped by the cell wall material, there are probably nucleases not only in the synergids but also in the pollen tube; there is no transport system known. Because of the attractiveness of this approach it is still worthy of rigorous testing.

10. Liposome fusion with tissues and protoplasts[23–25] (Fig. 10). *Transgenic plants have been recovered from protoplasts but not from tissues and cells; very little potential with cells.* Fusion of DNA-containing liposomes with protoplasts is an established method for the production of transgenic plants. It has, however, no obvious advantage over direct

gene transfer (see section 12) and electroporation (see section 21). DNA-containing liposomes have also been applied to various tissues, cell cultures and pollen tubes, with the rationale that liposomes might help transport the DNA via plasmodesmata or directly across the cell wall. It has been shown that liposomes can carry small dye molecules into cells within tissues via fusion with the plasmalemma. There is, however, no proof for transport and integration of marker genes. As plasmodesmata are sealed off immediately upon wounding this route is not open even for small liposomes; impregnation of the cell wall with phospholipids does not seem to change its barrier function.

11. Liposome injection[26] (Fig. 11). *Thus far no transgenic tissue recovered.* Microinjection of DNA has yielded transgenic chimeras (see section 17). Microinjection into differentiated cells can easily deposite the DNA into the vacuole, where it is degraded. W. Lucas et al.[26] had the elegant idea to exploit the vacuole. Microinjection of liposomes into the vacuole leads to fusion with the tonoplast thus releasing the content of the liposome into the cytoplasm, as demonstrated with cytoplasm-activated fluorescent dyes. Activity of injected DNA has still to be shown. This method, though elegant, has probably no advantage over straight forward microinjection especially for the production of transgenic cereals. Cereals regenerate only from meristematic cells that do not have large vacuoles.

12. Protoplasts and direct gene transfer[27–35] (Fig. 12). *The only method that has thus far yielded transgenic cereals; still problematic because plant regeneration from protoplasts is difficult to achieve.* Protoplasts efficiently take up DNA if treated with polyethyleneglycol (PEG) and/or electroporation. The protoplast isolation procedure probably shifts potentially competent cells to the competent state. If protoplast populations are available that contain competent cells, exogenous DNA is easily integrated via nonhomologous recombination. Also homologous recombination occurs, but at a far lower frequency. When protoplasts are transformed that are also competent for regeneration, transgenic plants can be recovered that stably contain, express and inherit the foreign gene. Protoplasts isolated from cereal tissues do not contain cells competent for regeneration (see section 13). Competent protoplasts have, so far, been isolated only from embryogenic suspensions established from immature tissues (scutellum, leaf base, anther). Standard direct gene transfer procedures with protoplasts from embryogenic suspensions has led to the regeneration of transgenic rice (*Oryza sativa* var japonica and *indica*) and maize (*Zea mays*) and will yield transgenic plants from other cereal species as soon as routine and efficient plant regeneration from protoplasts is established. This is, however, likely to be a problem for some years because, so far, the establishment of the appropriate cell cultures is an art that also depends upon parameters beyond experimental control. It is, therefore, difficult to envisage that this approach will ever serve as a solid basis for the required routine and efficient procedure for gene transfer into any desired species and variety.

13. Protoplasts from cereal plants[36] (Fig. 13). *No transgenic clones have been recovered; no potential to date.* As the establishment of appropriate embryogenic suspensions is a delicate and often unpredictable process, it would be of great advantage if protoplasts isolated directly from differentiated tissues could be cultured. However, this approach appears, to date, rather hopeless because differentiated cereal tissues do not express the wound response and obviously do not contain cells competent for regeneration. DNA uptake is no problem, as can be shown easily with transient expression assays. If integration occurs it has no consequences, because protoplasts do not proliferate. Although intensive experimentation has failed, so far, to induce proliferation, I would encourage further attempts.

14. Microlaser[37] (Fig. 14). *No transgenic tissue produced; not much potential.* A microlaser beam focussed into the light path of a microscope can be used to burn holes into cell wall and membranes. It was hoped that incubation of perforated cells in DNA solutions could serve as a basis for a vector-independent gene transfer method into walled plant cells. There are no conclusive data available on DNA uptake and there are problems with adsorption of exogenous DNA to cell wall material, even before it could be taken up. As microinjection and biolistics definitely transfers DNA into walled plant cells (see sections 16 and 17), the microlaser would offer advantages only in very specific cases where those techniques were not applicable.

15. Electrophoresis into tissues[38] (Fig. 15). *There is no proof for integrative transformation; judgement of the potential requires further experimentation.* Ahokas[38] tried to electrophorese DNA across the shoot meristem of barley seeds. His experiments yielded indicative evidence in the form of radioactively-labeled cell walls, positive GUS assays, and a protein on SDS-PAGE with *E. coli* GUS mobility. So far, all the data can also be interpreted as artifacts. It might, however, be worthwhile to test the potential of the idea with an experimental system that can give clearcut answers.

16. Biolistics or particle gun[39–42] (Figs. 16 and 17). *Thus far, no transgenic offspring produced in cereals. The method does, however, have good potential and is excellent for testing gene expression in transient systems.* Acceleration of heavy particles covered with DNA can be used to transport genes into plant cells and tissues. This technique caused some excitement because it was believed for some time that it would solve all gene transfer problems. Since then it has become possible to discuss "biolistics" on a more realistic basis. Transgenic plants have been produced in soybean and tobacco, and others will follow. The method has advantages that do give it the potential for general applicability: (1) it is easy to handle; (2) one shot can lead to multiple hits; (3) cells survive the intrusion of (one?) particle; (4) the genes coated onto the particle resume biological activity; (5) target cells can be as different as pollen, cell cultures, plant organs, and meristems; (6) particles also reach deeper cell layers. Thus the method provides a biological vector-independent DNA delivery system into a great variety of cells. Why then with all these advantages have no transgenic cereals been produced? As numerous particle gun machines are in use in many research groups, and many scientists have been shooting marker genes into many different experimental materials, we must assume that there are inherent problems. One of the technical problems for which improvement can be foreseen is the low frequency of transient and integrative events. As long as transient events occur at frequencies of ca. 10^{-4} and integrative events at ca. 10^{-8}, large numbers must be produced to hit one of the rare competent cells. Great effort and careful optimization of the parameters of the technique seem to improve the situation considerably. There remains the problem that the particles have to hit competent cells, which are very rare in cereals as long as embryogenic cell cultures are not available. As protoplasts isolated from embryogenic cell cultures definitely contain competent cells, and as these protoplasts are relatively easy to transform, direct gene transfer into protoplasts should be the method of choice

for those cases where such cultures have been established. The route via shoot meristems (in analogy to the successful soybean case) is considerably more difficult for cereals, because (a) the meristem is far better protected and more difficult to expose, (b) regeneration via multiple adventitious shoots from the tissue below the meristem is far less efficient, and (c) it is still an open question whether biolistics can transform meristematic cells. This is not unimportant for cereals because in cereals regeneration requires meristematic cells.

17. Microinjection[43,44] (Figs. 18 and 19). *No transgenic offspring have been recovered so far in cereals. The method does, however, have good potential.* Microinjection uses microcapillaries and microscopic devices to deliver DNA into defined cells in such a way that the injected cell survives and can proliferate. This technique has produced transgenic clones from protoplasts (where transformation via direct gene transfer is easier) and transgenic chimeras from microspore-derived proembryos in oilseed rape. As with biolistics, microinjection definitely delivers DNA into cells. In comparison with biolistics, microinjection has disadvantages (only one cell receives DNA per injection; the handling requires more skill and instrumentation). It also has advantages: (1) The quantity of DNA delivered can be optimised; (2) the experimenter can decide into which cell to deliver the DNA; (3) delivery is precise and predictable, even into the cell nucleus, and is under visual control; (4) small structures (e.g. microspores and few-celled proembryos, which are not available in the large quantities required for a biolistic experiment, can be targeted; (5) defined microinjected units can be micro-cultured; (6) if zygotic proembryos can be regenerated, microinjection would offer an approach to transformation that would be open for every species and variety having sexual propagation. *If competent cells could be visually identified,* no other technique could compete with microinjection. On the assumption that few-celled zygotic proembryos contain competent cells our group has established plant regeneration from isolated zygotic proembryos of maize *(Zea mays),* wheat *(Triticum aestivum),* rice *(Oryza sativa),* barley *(Hordeum vulgare),* soybean *(Glycin max),* cotton (*Gossypium* hybrid), sunflower *(Helianthus annuus),* tobacco *(Nicotiana tabacum),* and *Arabidopsis.* Following multiple microinjections with marker genes, putative primary transgenic chimeras and sexual offspring have been analysed (G. Neuhaus, G. Spangenberg, S. K. Datta. pers.comm.). So far, we have indicative evidence for putative transgenic chimeras. As we have no proof yet for transgenic offspring, these may be artifacts. Gene transfer into structures consisting of more than one cell can only produce transgenic chimeras (independent of the technique used), and transgenic offspring can only be expected if the transgenic sector contributes to the floral meristems. Therefore, two interpretations are possible to date (as for biolistics): (a) larger experiments will increase the chance for transmission to the offspring, or (b) meristematic cells are not competent for integrative transformation.

18. Zygotic proembryos and Agrobacterium (Fig. 20). To test the second hypothesis ((b) above) an experiment has been performed using a well-established and independent gene transfer system: pre-induced *Agrobacterium* carrying a proven resistance gene (35S-NPTII-35S) was used in an attempt to infect zygotic proembryos of *Nicotiana tabacum* var SR1 (a well-documented host plant for *Agrobacterium*-mediated gene transfer). The fact that transgenic tissue could not be detected, either in the regenerated tobacco plants or in the sexual offspring, unfortunately adds some weight to the second hypothesis (G. Neuhaus, A. Matzke, M. Matzke; pers. comm.).

19. Macroinjection[45] (Fig. 21). *There is no proof for the recovery of transgenic plants; probably no potential.* Use of injection needles with diameters greater than cell diameters leads to destruction of those cells into which DNA is delivered. DNA integration would require that the DNA moves into wound-adjacent cells and, therefore, all problems discussed in sections 7 and 8 apply. The most exciting data so far were reported in an experiment where a marker gene was injected into the stem below the floral meristem of rye *(Secale cerealse).* Hybridization to the marker gene and enzyme assays with selected sexual offspring yielded strong indicative evidence. Unfortunately, it has so far not been possible either to reproduce these data in several large-scale experiments with other cereals or to establish proof with the original material. It would be very difficult to understand how the DNA could reach the sporogenic cells in this experimental design, as DNA would not only have to reach neighbouring cells but would have to travel across many layers of cells.

20. Pollen transformation[46–49] (Fig. 22). *No transgenic plants have been produced; probably no potential.* This approach goes back to the early seventies and is based on the hope that DNA could be taken up into germinating pollen and either integrate into the sperm nucleus or reach the zygote with the pollen tube. Indeed, if this would function, this would be the ideal method for gene transfer into plants. Pollination with pollen germinated in the presence of DNA has yielded surprising results that could be interpreted as indicative evidence for gene transfer. In no case, however, has proof been provided nor could the transfer of a phenotype be shown to be caused by the transfer of a corresponding gene. As numerous large-scale experiments in experienced laboratories with defined marker genes have only given clearly negative results, it seems justified to conclude that this approach is not a very promising one. It is also understandable as not only the cell wall presents a problem, but also external and internal nucleases. The latter problem may be overcome with the technique of *in vitro* maturation[50] where immature microspores are treated with DNA, matured to pollen and then used for pollination, but here again there is no proof yet for transformation.

21. Electroporation[50–52] (Fig. 23). *No transgenic clones have been produced when applied to cells and tissues; not much potential with walled cells; routine method for gene transfer to protoplasts.* Protoplasts can be transformed with polyethyleneglycol (PEG), PEG + electroporation, electroporation alone, microinjection, and *Agrobacterium.* For protoplast systems, electroporation is but one of several modifications of direct gene transfer. Since in numerous important cases plants can be regenerated from cell cultures and tissue explants but not from protoplasts, it has been important to test whether electroporation can transfer genes into walled cells. This does not appear to be the case.

SUMMARY

Of numerous approaches to cereal transformation, so far only direct gene transfer into protoplasts has been successful. Considering the biological parameters outlined in section 1, it is not surprising that most approaches have not worked, and probably can not work. Biolistics, microinjection and *Agrobacterium* (?) have potential for a breakthrough. The fact, however, that careful and large-scale experiments with biolistic devices and microinjection into meristems and microspore-derived, zygotic and somatic proembryos have not yet yielded proof for the recovery of transgenic offspring may point to a biological problem that has not been considered in previous experiments, and which may also be the cause for the failure of

Agrobacterium-mediated transformation. Accumulated experience of gene transfer experiments with plants is in agreement with the hypothesis that meristematic (embryonic) cells can not be transformed. I do not know of any experiment that would disprove this hypothesis. It is, therefore, an important challenge for those interested in gene transfer to cereals to find out whether this is true, and if so, what prevents integrative transformation in these cells. Answers to these questions will be important independent of whether biolistics, microinjection, or *Agrobacterium* is used for gene transfer. It will, probably, be crucial for cereal transformation to solve this problem because it is not easy to regenerate cereals from non-meristematic cells. Although transgenic cereals can be regenerated from protoplasts in rice, and one has reason to hope that this will also be possible from other cereals, it would be unfortunate if gene technology with cereals has to rely on this tedious, unpredictable and unreliable method. If we can solve the problem of integrative transformation in zygotic proembryos, we can hope to have a method that can transfer genes into every variety of every plant species.

Acknowledgments

I am grateful to P. J. King, (SMI, Basel) for his help with the English version of this manuscript, and my numerous coworkers here in Zurich for their always helpful, always critical discussions.

Note added in proof: We have recently established what we believe is proof of the recovery of transgenic offspring of *Indica*-type rice. (Datta, S. K., Peterhans, A., Datta, K., Potrykus, I., 1990. Bio/Technology, submitted.

References

1. Gasser, C. S. and Fraley, R. T. 1989. Genetically engineering plants for crop improvement. Science **244**:1293–1299.
2. Kahl, G. 1982. Molecular biology of wound healing: the conditioning phenomenon. *In:* Molecular Biology of Plant Tumors. G. Kahl, J. Schell (Eds.). p. 211–267, Academic Press, NY.
3. Vasil, I. K. 1987. Developing cell and tissue culture systems for the improvement of cereal and grass crops. J. Plant Physiol **128**:193–218.
4. Hooykaas, P. J. J. 1989. Transformation of plant cells via *Agrobacterium*. Plant Mol. Biol. **13**:327–336.
5. Klee, H. J. and Rogers, S. G. 1989. Plant gene vectors and transformation: Plant transformation systems based on the use of *Agrobacterium tumefaciens*, p. 2–25. *In:* Cell Culture and Somatic Cell Genetics. Vol. 6. Molecular Biology of Plant Nuclear Genes. Schell, J. and Vasil, I. K. (Eds.), Academic Press, San Diego.
6. Binns, A. N. 1990. Agrobacterium-mediated gene delivery and the biology of host range limitations. Physiol. Plant. **78(1)**, *In press.*
7. Tepfer, D. 1990. Genetic transformation using *Agrobacterium rhizogenes*. Physiol. Plant. **78(1)**, *In press.*
8. Graves, A. C. F. and Goldmann, S. L. 1986. The transformation of *Zea mays* seedlings via the *Agrobacterium tumefaciens*. Plant Mol. Biol. **7**:43–50.
9. Schäfer, W., Görz, A., and Kahl, G. 1987. T-DNA integration and expression in a monocot crop plant after induction of *Agrobacterium*. Nature **327**:529–532.
10. King, P. J., Potrykus, I., and Thomas, E. 1978. *In vitro* genetics of cereals: problems and perspectives. Physiol. Vég. **16**:381–399.
11. Grimsley, N. H., 1990. Agroinfection. Physiol. Plant. **78(1)**, *In press.*
12. Gardner, R., Chonoles, K., and Owens, R. 1986. Potato spindle tuber viroid infections mediated by the Ti-plasmid of *Agrobacterium tumefaciens*. Plant Mol. Biol. **6**:221–228.
13. Grimsley, N. H., Hohn, T., Davies, J. W., Hohn, B. 1987. *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. Nature **325**:177–179.
14. Fütterer, J., Bonneville, J. M., Hohn, T. 1990. Cauliflower mosaic virus as a gene expression vector for plants. Physiol. Plant. **78(1)**, *In press.*
15. Brisson, N., Paszkowski, J., Penswick, J. R., Gronenborn, B., Potrykus, I., and Hohn, T. 1984. Expression of a bacterial gene in plants by using a viral vector. Nature **310**:511–514.
16. Ahlquist, P. and Pacha, R. F. 1990. Gene amplification and expression by RNA viruses and potential for further application to plant gene transfer. Physiol. Plant. **78(1)**, *In press.*
17. Ahlquist, P., French, R., Bujarski, J. J. 1987. Molecular studies of brome mosaic virus using infectious transcripts from cloned cDNA. Adv. Virus Res. **32**:214–242.
18. Töpfer, R., Gronenborn, B., Schaefer, S., Schell, J., and Steinbiss, H. H. 1990. Expression of engineered wheat dwarf virus in seed derived embryos. Physiol. Plant. **78(1)**, *In press.*
19. Töpfer, R., Gronenborn, B., Schell, J., and Steinbiss, H. H. 1989. Uptake and transient expression of chimeric genes in seed-derived embryos. The Plant Cell **1**:133–139.
20. Ledoux, L. and Huart, R. 1969. Fate of exogenous bacterial desoxyribonucleic acids in barley seedlings. J. Mol. Biol. **43**:243–262.
21. Ledoux, L., Huart, R., and Jacobs, M. 1974. DNA-mediated genetic correction of thiaminless *Arabidopsis thaliana*. Nature **249**:17–21.
22. Luo, Z. and Wu, R. 1988. A simple method for transformation of rice via the pollen-tube pathway. Plant Mol. Biol. Reporter **6**:165–174.
23. Caboche, M. 1990. Liposome-mediated transfer of nucleic acids in plant protoplasts. Physiol. Plant. **78(1)**, *In press.*
24. Gad, A. E., Rosenberg, N., and Altmann, A. 1990. Liposome-mediated gene delivery into plant cells. Physiol. Plant. **78(1)**, *In press.*
25. Ahokas, H. 1987. Transfection by DNA-associated liposomes evidenced at pea pollination. Hereditas **106**:129–138.
26. Lucas, W. J., Lansing, A., deWet, J. R., and Walbot, V. 1990. Introduction of foreign DNA into walled plant cells via liposomes injected into the vacuole: a preliminary study. Physiol. Plant. **78(1)**, *In press.*
27. Paszkowski, J., Saul, M. W., and Potrykus, I. 1989. Plant gene vectors and genetic transformation: DNA-mediated direct gene transfer to plants, p. 52–68. *In:* Cell Culture and Somatic Cell Genetics of Plants. Vol. 6. Molecular Biology of Plant Nuclear Genes. Schell, J. and Vasil, I. K. (Eds.). Academic Press, San Diego.
28. Davey, M. R., Rech, E. L., and Mulligan, B. J. 1989. Direct DNA transfer to plant cells. Plant Mol. Biol. **13**:273–285.
29. Potrykus, I., Shillito, R. D., Saul, M. W., and Paszkowski, J. 1985. Direct gene transfer. State of the art and future potential. Plant Mol. Biol. Rep. **3**:117–128.
30. Paszkowski, J., Baur, M., Bogucki, A., and Potrykus, I. 1988. Gene targeting in plants. EMBO J. **7**:4021–4027.
31. Shimamoto, K., Terada, R. Izawa, T., and Fujimoto, H. 1989. Fertile rice plants regenerated from transformed protoplasts. Nature **338**:274–276.
32. Rhodes, C. A., Pierce, D. A., Mettler, I. J., Mascarenhas, D., and Detmer, J. J. 1989. Genetically transformed maize plants from protoplasts. Science **240**:204–207.
33. Vasil, I. K. 1988. Progress in the regeneration and genetic manipulation of cereal crops. Bio/Technology **6**:397–401.
34. Shillito, R. D., Carswell, G. K., Johnson, C. M., DiMaio, S., and Harms, C. T. 1989. Regeneration of fertile plants from protoplasts of elite inbred maize. Bio/Technology **7**:581–587.
35. Prioli, L. M. and Söndahl, M. R. 1989. Plant regeneration and recovery of fertile plants from protoplasts of maize. Bio/Technology **7**:589–594.
36. Potrykus, I. 1980. The old problem of protoplast culture: cereals, p. 245–254. *In:* Advances in protoplast research. Ferency, L., Farkas, G. L. (Eds.). Pergamon Press, Oxford.
37. Weber, G., Monajembashi, S., Greulich, K. O., and Wolfrum, J. 1990. Genetic changes induced in higher plant cells by a laser microbeam. Physiol. Plant. **78(1)**, *In press.*
38. Ahokas, H. 1989. Transfection of germinating barley seed electrophoretically with exogenous DNA. Theor. Appl. Genet. **77**:469–472.
39. Sanford, J. C. 1990. Biolistic plant transformation. Physiol. Plant. **78(1)**, *In press.*
40. Klein, T. M., Fromm, M. E., Gradziel, T., and Sanford, J. C. 1988. Factors influencing gene delivery into *Zea mays* cells by high-velocity microprojectiles. Bio/Technology **6**:559–536.
41. McCabe, D. E., Swain, W. E., Martinell, B. J., and Christou, P. 1988. Stable transformation of soybean *(Glycine max)* by particle acceleration. Bio/Technology **6**:923–926.
42. Christou, P., McCabe, D. E., and Swain, W. F. 1988. Stable transformation of soybean callus by DNA-coated gold particles. Plant Physiol. **87**:671–674.
43. Neuhaus, G., Sopangenberg, G., Mittelsten-Scheid, O., and Schweiger, H. G. 1987. Transgenic rapeseed plants obtained by microinjection of DNA into microspore-derived proembryoids. Theor. Appl. Genet. **75**:30–36.
44. Neuhaus, G. and Spangtenberg, G. 1990. Plant transformation by microinjection technique. Physiol. Plant. **78(1)**, *In press.*
45. De la Pena, A., Lörz, H., and Schell, J. 1987. Transgenic plants obtained by injecting DNA into young floral tillers. Nature **325**:274–276.
46. Hess, D. 1987. Pollen based techniques in genetic manipulation. Int. Rev. Cytol. **107**:169–190.
47. DeWet, J. M. J., Bergquist, R. R., Harlan, J. R., Brink, D. E., Cohan, C. E., Newell, C. A., and DeWet, A. E. 1985. Exogenous DNA transfer in maize *(Zea mays)* using DNA-treated pollen, p. 197–209. *In:* Experimental Manipulation of Ovule Tissue. Chapman, G. P., Mantell, S. H., Daniels, W. (Eds.). Longman, London.
48. Ohta, Y. 1986. High efficiency genetic transformation of maize by a mixture of pollen and exogenous DNA. Proc. Natl. Acad. Sci. USA **83**:715–719.
49. Alwen, A., Eller, N., Kastler, M., Benito-Moreno, R. M., and Heberle-Bors, E. 1990. Potential of *in vitro* pollen maturation for gene transfer. Physiol. Plant. **78(1)**, *In press.*
50. Fromm, M. E., Taylor, L. P., and Walbot, V. 1986. Stable transformation of maize after electroporation. Nature **319**:791–793.
51. Shillito, R. D., Saul, M. W., Paszkowski, J., Müller, M., and Potrykus, I. 1985. High frequency direct gene transfer to plants. Bio/Technology **3**:1099–1103.
52. Lindsay, K. and Jones, M. K. G. 1990. Electroporation of cells. Physiol. Plant. **78(1)**, *In press.*
53. Potrykus, I. 1990. Gene transfer to plants: assessment and perspectives. Physiol. Plant **78(1)**, *In press.*

Transgenic plants as vaccine production systems

Hugh S. Mason and Charles J. Arntzen

Transgenic plants that express foreign proteins with industrial or pharmaceutical value represent an economical alternative to fermentation-based production systems. Specific vaccines have been produced in plants as a result of the transient or stable expression of foreign genes. It has recently been shown that genes encoding antigens of bacterial and viral pathogens can be expressed in plants in a form in which they retain native immunogenic properties. Transgenic potato tubers expressing a bacterial antigen stimulated humoral and mucosal immune responses when they were provided as food. These results provide 'proof of concept' for the use of plants as a vehicle to produce vaccines.

The development of genetic transformation technology for plants has facilitated the study of plant gene expression, and has resulted in great progress toward the genetic design of plants with enhanced production traits (such as herbicide, insect and disease resistance). Recently, several academic and industrial laboratories have begun experimenting with transgenic plants as novel manufacturing systems. We introduced the concept of vaccine production in transgenic plants in 1992 (Ref. 1). This effort was stimulated by interest in evaluating the capacity of plants to produce different classes of proteins of pharmaceutical value, and because of the practical need for new technology for the production and delivery of inexpensive vaccines, especially in the developing world. Candidate vaccines should be a useful test system for evaluating the capacity of transgenic plants to produce pharmaceutically active proteins, because the immune system would amplify the biological response to even relatively low levels of foreign protein.

On a more applied level, the announcement of the Children's Vaccine Initiative[2] documented the need for new vaccine technology to combat infectious diseases. We hypothesized that plants could be a useful system for producing vaccines, because large amounts of antigen could be produced at a relatively low cost, using agriculture instead of sophisticated and expensive cell culture-based expression systems. In this review, we will discuss the progress that has been made by several groups in what is now an expanding area of vaccine research that utilizes transgenic plants.

Oral vaccines and mucosal immune responses

Many infectious agents colonize or invade epithelial membranes; these include bacteria and viruses that are transmitted in contaminated food or water or by sexual contact. Vaccines that are effective against these infections must stimulate the mucosal immune system to produce secretory IgA (S-IgA) at mucosal surfaces such as the gut and respiratory epithelia[3,4]. In general, a mucosal immune response is more effectively achieved by oral, rather than parenteral, antigen delivery. Several particulate antigens have proven to be effective oral immunogens, including live and killed microorganisms. By comparison with parenteral immunizations, oral immunization using subunit or soluble antigens is often inefficient at stimulating an immune response, and requires larger amounts (mg versus μg) of antigen[5].

Subunit vaccines based upon recombinant cell-culture expression systems are feasible but, for commercial-scale production, these systems require fermentation technology and stringent purification protocols so that sufficient amounts of recombinant protein can be obtained for oral delivery. Even with technological improvements, fermentation-based subunit vaccine production may be a prohibitively expensive technology for developing countries where novel oral vaccines are urgently needed. Transgenic plants that express antigens in their edible tissue might be used as an inexpensive oral-vaccine production and delivery system[6]; therefore, immunization might be possible simply through consumption of an 'edible vaccine'.

The choice of which antigens to use in the initial studies has been strongly influenced by the desire to determine if transgenic plant materials containing foreign antigens will result in oral immunization and stimulate a mucosal immune response. Because of the need in the developing world for new oral-vaccine technology against diarrheal diseases[2], antigens from enteric pathogens have been the early targets for plant-based expression. In addition, effort has focused on the

H. S. Mason and C. J. Arntzen are at the Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14853-1801, USA.

production of antigens that assemble into ordered structures, such as virus-like particles (VLP), with the hope that they will be more resistant to digestion, more likely to reach the gut-associated lymphoid tissue (GALT) and, thus, more likely to be perceived as a foreign antigen once it reaches the GALT (Refs 3,4).

The choice of plant system for testing recombinant antigen production was initially driven by convenience and the need to evaluate genetic constructs quickly. For this reason, tobacco plants were generally utilized[1,7] but, because of the high levels of toxic alkaloids in the leaves, studies on animals feeding on tobacco are not practical without substantial purification of tobacco-derived antigens. Therefore, potato has been used as the plant of choice in several of the studies reported below[6,7]. This decision was based upon the observation that mice would accept raw potato tubers in lieu of laboratory feed, and because tubers can be generated for feeding studies within a few months of the transformation process[8]. With the purpose of acquiring a delivery system for humans, we have recently developed a genetic transformation system for banana[9], and are investigating the requirements for expressing abundant proteins in banana fruit. Crops used for animal feed, such as alfalfa, grains and beans, are obvious choices for animal vaccines, although the difficulty of generating uniform transgenic samples of these species has prohibited their use in most early studies.

Systems for expressing foreign proteins in plants

Two different strategies for transgene expression (for candidate vaccine production) in plants have been evaluated (Fig. 1). These involve either: (1) stable genomic integration, with foreign DNA introduced either by *Agrobacterium* T-DNA vectors or by direct means (including microprojectile bombardment); or (2) transient expression using viral vectors. Stable expression affords the advantage of the subsequent generation of large numbers of transgenic plants, either by vegetative or sexual means, and the opportunity to introduce more than one gene for possible multi-component vaccine production. In addition, judicious choice of genetic regulatory elements allows organ and tissue-specific expression of foreign antigens. Transient expression is less easy to initiate, because the viral vector must be inoculated into individual host plants; however, greater yields of foreign protein can usually be recovered. Published reports on both types of expression system are summarized below.

Stable genomic transformation using genes encoding foreign antigens

Streptococcus mutans *spaA protein*

The first report of the concept of using a plant expression system for the production of an edible vaccine appeared in a patent application published under the International Patent Cooperation Treaty[10]. It described a means to express a surface protein (spaA) from *S. mutans* in tobacco plants to a level of approximately 0.02% of the total leaf protein; the gene had been stably inserted by *Agrobacterium*-mediated transformation. Data were presented on the oral immunogenicity of spaA produced in *Escherichia coli*, which stimulated the production of S-IgA in saliva. No further reports of these studies have been published.

Hepatitis B surface antigen (HBsAg)

The expression of HBsAg at levels equal to 0.01% of total soluble protein in tobacco has been demonstrated[1]. The tobacco-derived recombinant HBsAg (rHBsAg) was recovered from leaf extracts as a VLP with an average size of 22 nm, which is important because the particle form of HBsAg is required for immunogenicity[11]. The plant-derived VLPs mimic the appearance of recombinant yeast-derived HBsAg particles[12], which is the material that is used in the currently available recombinant vaccine for hepatitis B (Recombivax®; distributed by Merck, Sharpe, and Dohme). In addition, the plant-derived material had similar buoyant density and antigenicity to human- and yeast-derived HBsAg, indicating faithful preservation of protein folding characteristics in the plant system[1].

A crude extract of rHBsAg from plants was used in parenteral immunization studies with mice[13]. The extract caused an immune response that was similar to the one achieved with Recombivax®, and included all the IgG subclasses, as well as IgM. Because T-cell-mediated immunity is essential for the prevention of hepatitis B, the fidelity of the T-cell epitope expressed by the tobacco-derived rHBsAg was investigated[13]. T cells isolated from the lymph nodes of mice, which were primed by parenteral immunization with the tobacco-derived rHBsAg, could be stimulated to proliferate *in vitro* by the tobacco-derived and yeast-derived rHBsAg. In total, the studies of rHBsAg from plants conclusively demonstrate that B- and T-cell epitopes of HBsAg are preserved when the antigen is expressed in transgenic plants, and that the recombinant antigen is produced as a VLP that mimics the currently available commercial vaccine.

E. coli *heat-labile enterotoxin B subunit and cholera-toxin B subunit*

In developing countries, diarrheal disease is an important cause of mortality, especially among children. Bacteria which cause diarrhea include *Vibrio cholerae* and the related enterotoxigenic *E. coli*. An oral vaccine composed of the cholera-toxin B subunit (CT-B) with killed *V. cholerae* cells gives protection against cholera and enterotoxigenic *E. coli* (Ref. 14). However, international health organizations have not distributed this vaccine because the cost of production of CT-B is too high for developing countries to afford.

The heat-labile enterotoxin (LT) from *E. coli* is a multimeric protein that is structurally, functionally and antigenically very similar to cholera toxin (CT). X-ray crystallography[15] has been used to determine that LT has one A subunit (LT-A), with an M_r of 27 kDa, and a pentamer of B subunits (LT-B), each of which has an M_r of 11.6 kDa. Specific binding of the nontoxic LT-B pentamer to the G_{M1} gangliosides present on

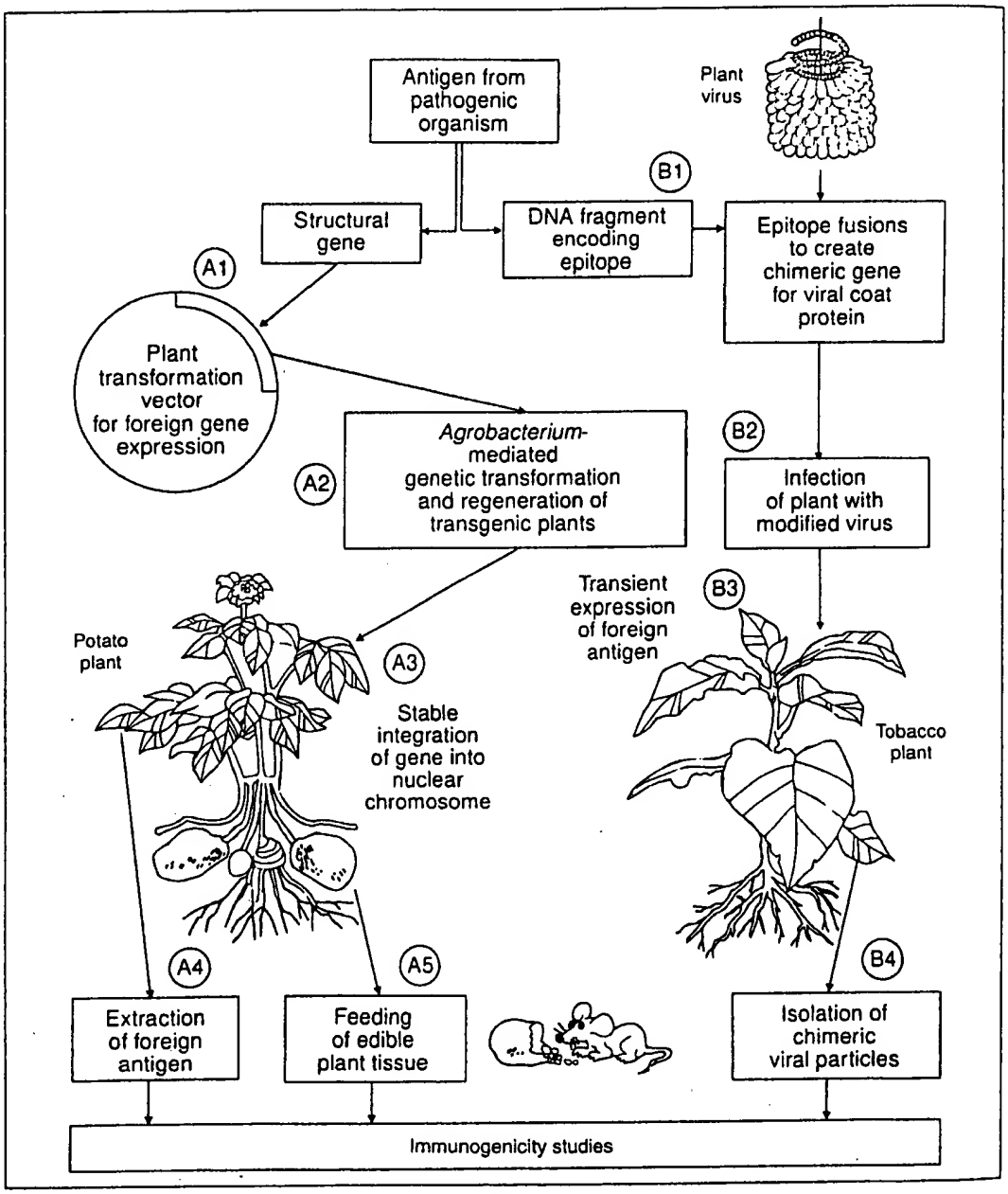


Figure 1

Strategies for the production of candidate vaccine antigens in plant tissues. Genes encoding antigens from pathogenic organisms (viruses, bacteria or parasites) that have been characterized and for which antibodies are available, can be handled in two ways. In one case, the entire structural gene is inserted into a plant transformation vector between 5′ and 3′ regulatory elements (**A1**); this will allow transcription and accumulation of the coding sequence in all, or selected, plant tissues. This vector is then used for the *Agrobacterium*-mediated transformation of plant cells (**A2**), or for stable integration of the expression cassette by other means, and regeneration of transgenic plants. The resulting plants contain the expression cassette stably integrated into the nuclear chromosomal DNA (**A3**), and can be used either for extraction and partial purification of the foreign antigen (**A4**), or for direct feeding of plant tissues (**A5**; in this case, a potato tuber) for assessment of immunogenicity. Alternatively, if epitopes within the antigen are identified, DNA fragments encoding these can be used (**B1**) to construct chimeric genes by fusion with a coat protein gene from a plant virus, e.g. tobacco mosaic virus (TMV) or cowpea mosaic virus (CPMV). The recombinant virus, or in the case of TMV and CPMV, even the viral RNA made *in vitro* from the plasmid clone, is then used to infect established plants (**B2**). Virus replication and systemic spread allow high-level transient expression of the chimeric coat protein in most plant tissues (**B3**). The viral particles, expressing the foreign epitope on their surfaces, are then purified (**B4**) and used for immunogenicity studies.

epithelial cell surfaces allows entry of the toxic LT-A subunit into cells. LT-B and CT-B are both potent oral immunogens[16].

LT-B has been expressed in transgenic plants, although the levels of expression were low[7]. The recombinant LT-B (rLT-B) produced in tobacco and potato showed enhanced accumulation when a C-terminal microsomal-retention sequence (Ser-Glu-Lys-Asp-Glu-Leu) was added. Sequestration of the rLT-B within microsomal vesicles may favor the association of individual subunits into the more stable pentameric form of LT-B. The tobacco-derived rLT-B appeared to be at least partially pentamerized, as judged by gel-permeation chromatography and its ability to bind gangliosides[7].

The oral immunogenicity of rLT-B was tested in mice and compared with bacterial LT-B (Ref. 7). When given orally to mice by gastric intubation, the plant-derived antigen stimulated humoral and mucosal immune responses, with titers comparable to the bacteria-derived LT-B. In addition, the antibodies produced against the tobacco-derived LT-B were able to neutralize LT activity, indicating the potential protective value of the immune response. The oral immunogenicity of unpurified rLT-B was also assessed by feeding raw transgenic potato tubers to mice. After only four feedings of 5 g tuber samples to mice, mucosal and serum antibodies were recovered. No immune response was observed in animals that were fed non-transformed tubers. This demonstrates that a food source containing a foreign antigen can induce oral immunization.

It should also be noted that CT and LT are excellent oral adjuvants, which stimulate immune responses against co-fed antigens[16,17]. Co-ordinate expression of A and B subunits to form the holotoxin in plants could enhance the vaccine value of other less immunogenic plant-expressed vaccine antigens expressed in the same tissues. This strategy is feasible because CT and LT function as adjuvants at concentrations well below those that cause diarrhea[16,17].

Transient expression of candidate vaccines using viral vectors

Using viral vectors for transient expression in plants represents a potentially useful means of producing high levels of recombinant antigens (Fig. 1). With tobacco mosaic virus (TMV), there are two ways to attain foreign-protein expression: (1) foreign-gene transcription, driven from a subgenomic promoter; and (2) fusion of foreign proteins or peptides with the capsid protein that normally coats the virus. The first strategy has been used to produce high-level expression of α-trichosanthin, an antiviral protein, in transfected plants[18], but the length of the foreign DNA insert that can be tolerated is undetermined. Capsid–protein fusions may be a better strategy, especially because the foreign protein is in particulate form (TMV virions), which is highly immunogenic[19]. Although the recombinant virus would need to be highly purified for parenteral administration, or partially purified for oral administration, this strategy may prove to be a cost-effective alternative to cell culture-based recombinant expression.

Malarial epitope fusions with TMV capsid protein

Turpen *et al.* have described a method for engineering the capsid protein of TMV as either internal or C-terminal fusions with peptides carrying epitopes derived from malarial sporozoites[19]. Both internal and C-terminal fusion constructs yielded high titers of genetically stable recombinant virus when used to infect tobacco plants. Antigenicity, measured by enzyme-linked immunosorbent assay (ELISA) and western blot, showed that the recombinant capsid proteins were recognized by the appropriate monoclonal anti-malarial antibodies.

Zona pellucida protein fusion with TMV capsid protein

The zona pellucida ZP3 protein of mammalian oocytes has been a target for immune contraception, and an epitope of 13 amino acids from murine ZP3 has now been expressed in plants as a fusion with TMV capsid protein[20]. The recombinant virus accumulated to high levels in infected plants, although systemic movement was somewhat slower than for the wild-type virus, and the viral particles were smaller. Mice immunized with the recombinant virus developed antibodies against ZP3 that were recruited to the zona pellucida; ovarian pathology was seen, but there was no observable effect on the fertility of the treated mice. Further work with extended epitopes or epitope mixtures may yield better results.

Cowpea mosaic virus capsid protein fusion

Cowpea mosaic virus (CPMV) has recently been developed as an expression system for the presentation of foreign peptides[21,22]. The advantages of CPMV are: high yield (1–2 g of virus per kg of host tissue), thermostability and ease of virus purification. CPMV is an icosahedral particle containing 60 copies each of large (L) and small (S) coat proteins. Earlier work showed that a foot and mouth disease virus epitope could be expressed as an S-protein fusion[21], but problems arose with the loss of the inserted RNA during serial passage. Modification of the chimeras resulted in genetically stable fusions, and the system has now been used to produce virus-expressing epitopes derived from human rhinovirus 14 and human immunodeficiency virus (HIV-1) that are immunogenic in test animals[22]. Furthermore, the antibodies raised against the CPMV–HIV chimera were able to neutralize three different strains of HIV-1 (Ref. 23). Inserts as long as 30 amino acids have been used in CPMV chimeras[21], although a particular sequence may have unpredictable effects on the systemic spread of the virus, thus compromising virus yield.

Future prospects

The demonstration that vaccine antigens can be produced in plants in their native, immunogenic forms opens exciting possibilities for the 'bio-pharming' of vaccines. If the antigens are orally active, food-based

'edible vaccines' could allow economical production and delivery in developing countries. Some engineering challenges remain, including maximizing the expression of the antigenic proteins, stabilizing the foreign protein during post-harvest storage in plant tissues, and enhancing the oral immunogenicity of some antigens.

Concerns about allergy and immune tolerance of orally applied antigens must be addressed as this technology develops, and will be resolved only by collaborative research efforts with medical specialists in these fields. It must be determined whether the levels of antigen required to induce the desired oral immune responses are less than the levels that could induce tolerance, as happens with proteins that occur as normal components of our food. If so, a delivery scheme must be developed to provide only the 'edible vaccine' as a medicinal product at the required dosage level and not as a routine food source. The specific oral response (immunogenicity versus tolerance) may be antigen specific and may, therefore, need to be considered on an individual vaccine basis. Research will be needed to determine which types of plants are most suitable for vaccine delivery; this must be coupled with the discovery of means to express the protein at the desired levels in the appropriate plant cells and tissues. Lastly, a thorough study of the relationship between immunogenic dose responses and antigen levels in various foods needs to be undertaken. In the short term, vaccines for animals are a more likely target for edible-vaccine technology than vaccines for humans, and studies in this area are likely to increase our understanding of the basic mechanisms, which can then be applied to the development of all vaccines.

References

1 Mason, H. S., Lam, D. M-K. and Arntzen, C. J. (1992) *Proc. Natl Acad. Sci. USA* 89, 11745–11749
2 Anonymous (1993) *The Children's Vaccine Initiative; Achieving the Vision* (Mitchell, V. S., Philipose, N. M. and Sanford, J. P., eds), National Academic Press
3 McGhee, J. R. and Kiyono, H. (1993) *Infect. Agents Disease* 2, 55–73
4 Castro, G. and Arntzen, C. J. (1993) *Am. J. Physiol.* 165, 599–610
5 de Aizpurua, H. J. and Russell-Jones, G. J. (1988) *J. Exp. Med.* 167, 440–451
6 Arntzen, C. J., Mason, H. S., Shi, J., Haq, T. A., Estes, M. K. and Clements, J. D. (1994) in *Vaccines '94: Modern Approaches to New Vaccines Including Prevention of AIDS* (Brown, F., Chanock, R. M., Ginsberg, H. S. and Lerner, R. A., eds), pp. 339–244, Cold Spring Harbor Laboratory Press
7 Haq, T. A., Mason, H. S., Clements, J. D. and Arntzen, C. J. (1995) *Science* 268, 714–716
8 Wenzler, H., Mignery, G., May, G. and Park, W. (1989) *Plant Sci.* 63, 79–85
9 May, G. D., Afza, R., Mason, H. S., Weicko, A., Novak, F. J. and Arntzen, C. J. (1995) *Bio/Technology* 13, 486–492
10 Curtiss, R. and Cardineau, G. A. (1990) *World Intellectual Property Organization*, PCT/US89/03799
11 Cabral, G. A. *et al.* (1978) *J. Gen. Virol.* 38, 339–350
12 Scolnick, E. M., McLean, A. A., West, D. J., McAleer, W. J., Miller, W. J. and Buynak, E. B. (1984) *J. Am. Med. Assoc.* 251, 2812–2815
13 Thanavala, Y., Yang, Y-F., Lyons, P., Mason, H. S. and Arntzen, C. J. (1995) *Proc. Natl Acad. Sci. USA* 92, 3358–3361
14 Clemens, J. D. *et al.* (1988) *J. Infect. Dis.* 158, 372–377
15 Sixma, T. K. *et al.* (1991) *Nature* 351, 371–377
16 Clements, J. D., Hartzog, N. M. and Lyon, F. L. (1988) *Vaccine* 6, 269–277
17 Holmgren, J., Lycke, N. and Czerkinsky, C. (1993) *Vaccine* 11, 1179–1184
18 Kumagai, M. H. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 427–430
19 Turpen, T. H., Reinl, S. J., Charoenvit, Y., Hoffman, S. L., Fallarme, V. and Grill, L. K. (1995) *Bio/Technology* 13, 53–57
20 Fitchen, J., Beachy, R. N. and Hein, M. B. *Vaccine* (in press)
21 Usha, R. *et al.* (1993) *Virology* 197, 366–374
22 Porta, C., Spall, V. E., Loveland, J., Johnson, J. E., Barker, P. J. and Lomonossoff, G. P. (1994) *Virology* 202, 949–955
23 McLain, L., Porta, C., Lomonossoff, G. P., Durrani, Z. and Dimmock, N. J. (1995) *AIDS Res. Hum. Retroviruses* 11, 327–334